

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:13:46 ; Search time 4994.97 Seconds
(without alignments)
1952.395 Million cell updates/sec

Title: US-09-214-124-1
Perfect score: 940
Sequence: 1 aauguggggggagcuccg.....ccucuccggccgggauggg 940

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 5187315402 residues
Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_bal: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl1: *
8: gb_pl2: *
9: gb_pr1: *
10: gb_pr2: *
11: gb_pr3: *
12: gb_ro: *
13: gb_sts: *
14: gb_sy: *
15: gb_un: *
16: em_fun: *
17: em_hum1: *
18: em_hum2: *
19: em_in: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_sy: *
29: em_un: *
30: em_v1: *
31: gb_htg1: *
32: gb_htg2: *
33: gb_in1: *
34: gb_in2: *
35: em_ba1: *
36: em_ba2: *
37: em_hum3: *
38: em_hum4: *
39: gb_pr4: *
40: gb_htg3: *
41: gb_htg4: *
42: gb_htg5: *
43: gb_htg6: *

44: gb_htg7: *
45: em_htg1: *
46: em_htg2: *
47: em_htg3: *
48: em_hum5: *
49: gb_pl3: *
50: gb_pr5: *
51: gb_htg8: *
52: gb_htg9: *
53: gb_htg10: *
54: gb_htg11: *
55: gb_htg12: *
56: gb_htg13: *
57: gb_htg14: *
58: gb_in3: *
59: gb_htg15: *
60: gb_htg16: *
61: gb_htg17: *
62: em_htg4: *
63: em_htg5: *
64: em_htg6: *
65: em_htg7: *
66: em_hum6: *
67: gb_htg18: *
68: gb_htg19: *
69: gb_htg20: *
70: gb_htg21: *
71: gb_htg22: *
72: gb_htg23: *
73: gb_htg24: *
74: gb_htg25: *
75: gb_htg26: *
76: gb_htg27: *
77: gb_htg28: *
78: gb_htg29: *
79: gb_htg30: *
80: gb_htg31: *
81: gb_v11: *
82: gb_v12: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	940	100.0	940	5	A83548	A83548 Sequence 1
2	825.2	87.8	4292	82	AF006065	AF006065 Fowlpox v
3	578	61.5	578	5	A83549	A83549 Sequence 2
4	529.8	56.4	545	81	S70398	S70398 (LTR, U3, R
5	512.4	54.5	583	81	S79845	S79845 (REV, LTR) I
6	496.6	52.8	891	81	REXXX1	V01204 Spleen necr
7	455.2	48.4	1005	81	ACRLTRL	M22223 Reticuloend
8	440.8	46.9	585	81	S82226	S82226 (RS region,
9	431.8	45.9	887	81	ACRLTRL2	M22224 Reticuloend
10	425.6	45.3	1530	81	RESNVX	V01200 Spleen necr
11	307.6	32.7	859	81	REXXX2	V01205 Spleen necr
12	160.4	17.1	1022	82	AF006066	AF006066 Fowlpox v
13	156	16.6	2640	82	AF006064	AF006064 Fowlpox v
14	156	16.6	288539	82	AF198100	AF198100 Fowlpox v
15	117.2	12.5	180	81	SNVLTR	X59450 spleen necr
16	113.6	12.1	160	4	CHKSNVIE2	M12248 Spleen necr
17	111.6	11.9	160	4	CHKSNVIB2	M12242 Spleen necr
18	111	11.8	160	4	CHKSNVID2	M12246 Spleen necr
19	111	11.8	160	4	CHKSNVIF2	M12250 Spleen necr
20	110.6	11.8	160	4	CHKSNVIA2	M12206 Spleen necr
21	110.6	11.8	160	4	CHKSNVIC2	M12244 Spleen necr
22	82.6	8.8	163	81	SNVPRTRMB	M12297 Spleen necr
23	73.2	7.8	100	81	SNVLTRI2	M19315 Spleen necr
24	60.2	6.4	209	81	SNVPRTRMA	M12208 Spleen necr

25	44.8	4.8	100	81	SNVTRIL	M19314 Spleen necr
26	44.8	4.8	290	13	AU025522	AU025522 Rattus no
27	42.6	4.5	63	4	CHKSNVIF1	M12249 Chicken DNA
28	42.6	4.5	69	4	CHKSNVIA1	M12205 Chicken DNA
29	42.6	4.5	69	4	CHKSNVIB1	M12241 Chicken DNA
30	42.6	4.5	69	4	CHKSNVID1	M12245 Chicken DNA
31	42.2	4.5	69	4	CHKSNVIC1	M12243 Chicken DNA
32	42.2	4.5	69	4	CHKSNVIE1	M12247 Chicken DNA
33	38.2	4.1	1143	4	BBU74351	U74351 Bulweria bu
34	37.6	4.0	378	13	AU026236	AU026236 Rattus no
35	37.6	4.0	7218	5	I66494	I66494 Sequence 14
36	37.2	4.0	113440	55	AC020935	AC020935 Homo sapi
37	37	3.9	70480	39	AC006973	AC006973 Homo sapi
38	37	3.9	173618	57	AC019045	AC019045 Homo sapi
39	36.4	3.9	1143	4	AF076053	AF076053 Fregetta
40	36.2	3.9	2041	12	MUSENBE1A	M88481 Mouse enzym
41	36.2	3.9	162384	51	AC007840	AC007840 Drosophill
42	36.2	3.9	195960	42	AC013569	AC013569 Homo sapi
43	36	3.8	16832	42	AC014255	AC014255 Drosophill
44	36	3.8	88866	41	AC010696	AC010696 Drosophill
45	36	3.8	114958	51	AC008356	AC008356 Drosophill

ALIGNMENTS

RESULT 1

LOCUS	A83548	940 bp	DNA	PAT	21-JAN-2000
DEFINITION	Sequence 1 from Patent WO9849334.				
ACCESSION	A83548				
VERSION	A83548.1	GI:6732807			

KEYWORDS

Reticuloendotheliosis virus.
Reticuloendotheliosis virus

REFERENCE

1 (bases 1 to 940)
Gabus-Darlux,C. and Darlix,J.
retroviruses; 3-Reticuloendotheliosis virus group.

AUTHORS

NOVEL INTERNAL RIBOSOME ENTRY SITE AND VECTOR CONTAINING SAME
Patent: WO 9849334-A 05-NOV-1998;

JOURNAL

GABUS DARLIX CAROLINE (FR); INST NAT SANTE RECH MED (FR)

FEATURES

location/Qualifiers

source

1..940

/organism="Reticuloendotheliosis virus"

/strain="TYPE A (REV-A)"

/isolate="LEADER 5' DE L'ARN GENOMIQUE REV-A"

/db_xref="taxon:11636"

BASE COUNT 190 a 225 c 260 g 265 t

ORIGIN

Query Match 100.0%; Score 940; DB 5; Length 940;
Best Local Similarity 71.8%; Pred. No. 1.2e-288;
Matches 675; Conservative 265; Mismatches 0; Indels 0; Gaps 0;

QY	1	aaugugggagggagcucggggggaauagcgucgucgucuaacugccauuagcuu	60
Db	1	AATGTGGAGGAGCTCCGGGGGAATAGCGCTCGCTTAAGTGCATATTAAGCTTCT	60
QY	61	guaucaugcugcugcucuuagccgccaugacuuaauuugcgauuauuu	120
Db	61	GTAATCATGCTTGCTTGCCTTAGCCGCATTTGATATATTTGCTGATATCATTT	120
QY	121	cucggaucgcauauuucggaucgcauagcagcagcucuaagcauuaaag	180
Db	121	CTCGGAATCGGCATCATTTCTCGGAATCGGCATCAAGAGCAGGCTCATTAAG	180
QY	181	gaaugucgugagggcgagcauagcagcagcagcagcagcagcagcagcag	240
Db	181	GAATGTGCTGAGGCGAGCATCAGACCACTTGCCGCATCCAATCAGCAAAACAG	240
QY	241	agaucgaacuaucugagccaaugguuuaagggcgagcagcuaucuccaauag	300

Db	241	AGATCGAATCATATCTAGTGAACCAATGTTGTAAGGCGAGATGCTATCTCCATGAGG	300
QY	301	gaaugucagcauacucgucgucgucgucgucgucgucgucgucgucgucgucg	360
Db	301	GAAATGTATGCAACATCTGCTCTGTAAGCGGCTATATAAGCCAGGTGCTCTTGC	360
QY	361	ucgugcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugc	420
Db	361	TCGGGCTCGCCGCTCTACATGTTGTGACGCGCGCCAGATTCGAATCTGATAA	420
QY	421	aguu	480
Db	421	AGTTTTTTCTCTATATCTCAGATTTGGCAGTAGAGAGATTTGTTGCTGCTAGG	480
QY	481	cuggcuaucugcugcugcugcugcugcugcugcugcugcugcugcugcugcugc	540
Db	481	CTGGCCTACTGGGTGGGTAGGGGTCCGGACTGATTCCTAGTATTTGATACACATTT	540
QY	541	gggggucgucgucgucgucgucgucgucgucgucgucgucgucgucgucgucg	600
Db	541	GGGGGCTCGTCCGGGATTCCTCCCATCGGAGAGTGCTTCTTCTCGAATCCG	600
QY	601	gcgcgcgucgucgucgucgucgucgucgucgucgucgucgucgucgucgucguc	660
Db	601	CGCCCGTAAGTACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	660
QY	661	cggagcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugc	720
Db	661	CGGACGCTGCGGGAAGCTCCACCTCCGCTCAGCAGGAGCGCCCTGATCTGAGCTCTG	720
QY	721	uguaucgucgucgucgucgucgucgucgucgucgucgucgucgucgucgucguc	780
Db	721	TGTTATCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	780
QY	781	uguu	840
Db	781	TGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	840
QY	841	gccacacgc	900
Db	841	GCCACACCGCGCGGCTTGGCAATATACTTTGGAGAGTCTTTTGGCTCCAGTGTCTCC	900
QY	901	guuugcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugcugc	940
Db	901	GTTTGTACTGCT	940

RESULT 2

AF006065

LOCUS AF006065 4292 bp DNA VRL 16-SEP-1997
DEFINITION Fowlpox virus S gag gene, complete cds, and protease/polymenase
(pro/pol) gene, partial cds.

ACCESSION AF006065

VERSION AF006065.1 GI:2393892

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

97428585

2 (bases 1 to 4292)

Hertig,C., Coupar,B.E., Gould,A.R. and Boyle,D.B.

Field and vaccine strains of fowlpox virus carry integrated

sequences from the avian retrovirus, reticuloendotheliosis virus

Virology 235 (2), 367-376 (1997)

Submitted (30-MAY-1997) Division of Animal Health, CSIRO, 5

Portarlington Road, Geelong, Victoria 3213, Australia

Location/Qualifiers

1..4292

source


```

/organism="Fowlpox virus"
/strain="S (standard vaccine strain)"
/db_xref="taxon:10261"
/note="standard vaccine strain from Cyanamid-webster Pty
Ltd, Castle Hill, New South Wales, Australia."
EcORI/PstI 4.3kb fragment from the 9.8kb PstI genome
fragment
integrated reticuloendotheliosis virus"

```

```

/note=
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC58238.1"
/db_xref="GI:2393893"
/translation="MATVARMYKTIINTGISCVLKSLIPDSYNEEYNIDDLIKRIKE
/translation="KSVTDSIVLYIENTNRYSISTSHDKNEPYEENGIIIMNNIECFYA
FLEISMORCFYSIKSVTDSIVLYIENTNRYSISTSHDKNEPYEENGIIIMNNIECFYA
FLEISMORCFYSIKSVTDSIVLYIENTNRYSISTSHDKNEPYEENGIIIMNNIECFYA
CLEGSCIVANVLGDROI SDNISSESGFLMDVNDHVIDIKYVGLFTIKVDAHVEYG
QNVIMEPEKNLFSQTNGPNFLYDITVQDRNVLLITSKYIYNLCDDKIYDIFELKYL
QNVIMEPEKNLFSQTNGPNFLYDITVQDRNVLLITSKYIYNLCDDKIYDIFELKYL
VDCNKLPMPLIPLSKYDFTFTLSVIXKSENVKTYLSKVHTSKSYINNDTSLPVAVKV
YGVTVTIR"
1812..3311
/gene="gag"
1812..3311
/gene="gag"

```

gene
CDS

[illegible]

	Query Match	Best Local Similarity	Mismatches	Score	Pred.	No. Indels	Gaps
OY	7	ggaggagcuccgsggggaauagcgucgucgcuaacugcccauuuagcuucguuanc	66	87.8%;	68.3%;	253;	DB 82; Length 4492;
Dd	908	GGAGGAGCTCCGGGGGAATAGCGCTGCCTACGTACCATTATTTGGCTGAT	18;				
OY	67	augcugucguucgcuuagccgcccauugacuauuauuucgucgaaucuuucgcga	126				
Dd	968	ATGCTTGCTGCCCTTAGCCGCCATTGTAATTGATTTGGCTGAT	1014				
OY	127	aucgagcauuucucgcgaaucgcaucaagaagcagcgucuaagaccuanaagaaug	186				
Dd	1015	-----ATCATTTCTCGGAATCGGCATCAGAGACGAGGCTCATAAACCAATAAAGAATG	1068				
OY	187	uucguugagcgagcgaucagaccacuuucgccaaccaucaacgagcaaacagagaucy	246				

Db	1069	TTTGTGAAGGCAAGCATCAGACCACCTTGACCAATCCATCAGCAACAAACACGAGATCG	1128
QY	247	aacuaucuaucugagcgcaauuguuuaaaggcgagauycuaucucccauagggagaaau	306
Db	1129	AACTATCATCTAGAGCCAAATGGTTGTAAGGGCAGATGCTATCTCCATATGAGGAAAT	1188
QY	307	gucaugcaacauccuguccuguaaagcgcuauuaaagccagugcaucucugcggg	366
Db	1189	GTCAATGCAACA-----TCCTGTAGCGGGCTATATAAAGCAGGTGCATCTTGTCTGGGG	1243
QY	367	ucgcccgcuaacacauuugugagcgcgcccgagauucgaaucugaaauaaguuuu	426
Db	1244	TCGCCGTCTACACATTTGTGTAGCGTCCGCCAGATTGCAATCTGTAATAAAG-CTT	1302
QY	427	uuucuucauauccucagauugcgagugagggaguuuugucugugugagcgcc	486
Db	1303	TTTCTCTATATCTCTCAGATTGGCAGTGAAGAGAGATTGTGTCTGTGTGTGGCTGGCC	1362
QY	487	uacugugugggguaggggucgcgagcuaaucgaaugaaauuucgaaacaauuggggc	546
Db	1363	TACTGGTGGGTAGGGATCCGGACTGAATCCGTAGTATTCGTACACATTTGGGGCC	1422
QY	547	ucgucgcggauucccccacucggcagaugcuaucuguuuucuuugaaucucgcgcgcg	606
Db	1423	TCGTCGGGATTCCTCCCATCCGAGAGAGGTGCTTCTTCTGCAACTCCGGCGCCG	1482
QY	607	guaauguaquacuugauuuuguaucucgcgagguuuugaggauccgagugcgagac	666
Db	1483	GTAAGTAAGTAATTGATTTGTGTACTCCGAGGGTTGGAGGTTCCGAGTGGCGGAC	1542
QY	667	gcugccggagagcuccacucgcgcucagcagggagcccgauucugagcucuguguan	726
Db	1543	GCTGCCGGGAGCTCCACCTCCGCTCAGCAGGGGACGCCCTGGTGTAGCTCTGTGTAT	786
QY	727	cugauuguuugugagcgcugucuccaagcgcgugaaauaauaagucguuuuguuug	1602
Db	1603	CTGATTTGTGTGAACCGTCTTAAGACGGTGTACTATAAGTCTGTGTGTGTG	1662
QY	787	uuuguuaccuuguguuugucguacauugucagcagcccgccugcgaaauuguguccac	846
Db	1663	TTTGTTAACCTGTGTGTGTCTGTCTGACAGCGCCCTGCGAATTGGTGTACCCAC	1722
QY	847	accgcgcgcguugcgaaauaauacuugagagucuuuugcuccagugucuccguuug	906
Db	1723	ACCGCGCGGCTTGGCAATAATACTTTGGAGAGTCTTTGGCTTCCAGTGTCTCCGTTGT	1782
QY	907	acucgucucucucucucucucucgcgcggggauggg	940
Db	1783	ACTGTCCTCTCTCCCTCTCTCCGGCGGATGGG	1816

	PAT	21-JAN-2000
JOURNAL TITLE	DNA	
REFERENCE AUTHORS	A83549	578 bp
SOURCE ORGANISM	Sequence 2 from patent WO9849334.	
KEYWORDS	A83549	
	A83549.1	GI:6732808
	Reticuloendotheliosis virus.	
	Reticuloendotheliosis virus	
	Reticuloendotheliosis viruses; Retroviridae; Mammalian type C	
	viruses; Retroid viruses; Reticuloendotheliosis virus group.	
	retroviruses; 3-Reticuloendotheliosis	
	1 (bases 1 to 578)	
	Gabus-Darlix.C. and Darlix.J. SITE AND VECTOR CONTAINING SAME	
	NOVEL INTERNAL RIBOSOME ENTRY	
	Patient: WO 9849334-A 05-NOV-1998;	
	GABUS DARLIX CAROLINE (FR); INST NAT SANTE RECH MED (FR)	
	Location/Qualifiers	
FEATURES	1..578	
source	/organism="Reticuloendotheliosis virus"	
	/strain="TYPE A (REV-A)"	
	/db_xref="taxon:11636"	
	/db_xref="taxon:11636"	
BASE COUNT	94 a	139 c
	139 g	173 t

BASE COUNT	ORIGIN	Score	DB	Length	Gaps
165 a	/db_xref="taxon:10390"	512.4	81	583	1
145 c	123 g	152	1		
149 t	1 others				

Query Match	70.1%	Mismatches
Similarity	147;	
Best Local		
Conservative	386;	
Matches		

```

CGGAGCUCGGGGGAUAGCGUGCUACUGCACAUAUAGUCUCC
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CGGCTTAACTGCATATTAGCTTC

```

504
120

[illegible][illegible][illegible]

383 GAATGTTGTGGAGGCGCCTCCTCCATGAGG

UY |||||
323 AGATCGAACTACATCATCAGGCGGCAUUAAGCCAAAGTCTCTGGC
Db |||||

263 GAAATGTCTATGCACATC
Db

203 :CGGGGTCGCCGCTCTCAAC...
 Db :CGGGGTCGCCGCTCTCTCGTGC...

Accession	Gene	Protein	Length	Score	E-value
U7	ATTTTTCCTTCTATAATCC	143	143	143	143
Db	ATTTTTCCTTCTATAATCC	143	143	143	143

[illegible]

23 TAGTTTTCAT 13

	6	DNA	Integrated	--
RESULT	891 bp			
REXXx1	necrosis virus (5' end)			
REXXx1	colicin m12296			
LOCUS				

ACCESSION
VERSION
KEYWORDS
V01204.1
Enjeen necrosis virus.
Enjeen necrosis virus
Retroviridae; Mammalian ty

ORGANISM
Viruses; Retroviruses; I-Mammalian life
retroviruses; 1-Mammalian life
bases 1 to Mutant, S. and Temin, H.M.
1 (bases 1 to Mutant, S. and Temin, H.M.)
1 (bases 1 to Mutant, S. and Temin, H.M.)

AUTHORS	TITLE
Sequence of elements transposable	550-554 (1980)
Nature	285 (5766),

source	FEATURES	MEDLINE
1. .891	/organism="spleen necrosis	Viru
at "ref="taxon:11836"		

```
source
1:1:0="cellular" 0:0:0
```

```

/organism="Spleen necrosis virus"
/db_xref="taxon:11836"
1..891
/organism="Gallus sp."
/ah_xref="taxon:9036"
source

```

```

/locus>891
26..>891
/organism="Gallus sp."
/proviral
...9036"
source

```

BASE COUNT	194 a	215 c	240 d
	NR 81;	Length 891;	121;

Query Match	Similarity	Pred. NO.	Mismatches	Conservative
52.8%	62.6%	192	69	158
Best Local	509			
Matches				

DG
99 GCCTGGCCACTATACCGGCATATTAGC-
.....:::|||||CCCCCATCAAGTTTCG-----CT 20

[illegible][illegible]

263 ACTGCCCATCATCATGACACCGCTGAGCTC

323 TGTAAAGGCAGATGCTACTACACCA
DB

[illegible][illegible]

501 ATTGGCAGTGAGAGGAGATTTC...
 Db ATTGGCAGTGAGAGGAGATTTC...

OY :|||||ATCCGAGTAC...
 DB :|||||uuucuuagacucugy|||:
 561 ATCCGACTGAATCCGGCGCCGCT
 :|||||tttcttcgactccggcgccgct
 :cgaacugcuacug--
 :cgaacugcuacug--

[illegible]

681 TTAGACTGTGATTTGGTACC
Db

[illegible][illegible]

859 Db

1

Pa

DB	401	G-CTTTTCTCTATATCCATGATTGGCAGTGAGAGAGATTGTTCTGTTGTTGGC	459
QY	482	uggcucacugggugggguagggggucgcgacugaaucguaguauuucgaaacaaauug	
Db	460	TT	
QY	542	ggggcucgucgcgggag 557	
Db	520	GGGGCTCGTCCGGCAT 535	
RESULT	8		
S82226			
LOCUS			
DEFINITION	S82226	585 bp	DNA
ACCESSION		(RS region, intervening sequence)	VRL
VERSION	S82226		12-FEB-1997
KEYWORDS	S82226.1	GI:1839499	
SOURCE			
ORGANISM	Gallid herpesvirus 2 RM1.		
REFERENCE	Gallid herpesvirus 2		
AUTHORS	1 (bases 1 to 585)		
TITLE	Jones, D., Brunovskis, P., Witter, R. and Kung, H.J.		
JOURNAL	Activation of U(S) genes by an integrated long terminal repeat in a		
MEDLINE	J. Virol. 70 (4), 2460-2467 (1996)		
REMARK	GenBank staff at the National Library of Medicine		
COMMENT	This sequence is identical to the one in GenBank		

```

Query Match
Best Local Similarity 46.9%; Score 440.8; DB 81; Length 585;
Matches 370; Conservative 136; Mismatches 17; Indels 25; Gaps 3;

QY 1 aaugugggagggaagcuccgggggggaauagcgucgucgcuaacugccauuagucuu 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 44 AATGTGGAGGAGGAGCTCCGGGGGAATAGCCCTGGCTGCTAACTGGCCATATTAGCTTCT 103
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 guaaucagucguucguccguuagccgcacauuguaucuguaauuuucguguaucuu 120
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 GTAATCATGCTTGTGCTTGCCTTAGCCGCCATTGTTGACTTGATATATTTCGCTGAT 156
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 -----ATCATTTCTCGGAATTCGGCATCAAGAGCAGGCTTCATAACCATATAAAG 204
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 gaauuguuucguugaggcgagcaucagaccacucgcgcauccaaucaagagcaaacag 240
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 205 GAAATGTTGTGGAAGGCAAGCATCAGACCACCTTGACCATCCAAATCAGAAACAACACG 264
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 agaucgaucuaucacugagcgcaauuguuuuaaaggcgagaugcuauccuccaauagag 300
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 265 AGATCGAACTATCATTACTGAGCCAAATGTGTCTAAAGGCGCATGTCTATCTCCAATGAGG 324
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 gaauuuguaucgcaacacuccugucuuagcggcuaauaagccaagugugcaucucugc 360
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 325 GAAAAATGTCATGCAACA-----TCCTGTAAGCGGCTATATTAAGCCAGGTGCATCTTGC 379
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 ucggggugcgcgucucuaacacauuuguuugacgcgcgccagaucgaucuguaauaa 420
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db	380	TCGGGGTCGCCGCTCCTACACATTGTTGTGACGTCGGGCCAGATTCCGATCTGTAATAA	439
OY	421	AGUUUUUUUCUUAUAUCUUCAGAUUGGCGAGUGAGGAGAUUUUGUUCGUGUGUAGG	480
Db	440	AG-CTTTTCTCTATATCCTCAGATTGGCAGTAGAGAGAGATTGTGTCGTGCTTG	498
OY	481	CUGGCCUAUCUGUGUGUGGUAAGGUGUCCGGACUGAAUCCGUAGUAUUUGCAACAACAUU	540
Db	499	CTGGCCCTACTGGGTGGGTAGGATCCGCACTGAATCCGTAATTTCCGTAACAACAGG	558
OY	541	GGGGGCGUC	548
Db	559	GGGAAGTC	566

RESULT	9				
ACRLTR2					
LOCUS	ACRLTR2	887 bp ss-RNA	VRL	27-APR-1993	
DEFINITION	Reticuloendotheliosis virus	713	proviral clone	RNA.	
ACCESSION	M22224	M10660	M16722		
VERSION	M22224.1	GI:209709			
KEYWORDS	long terminal repeat (LTR).				
SOURCE	Reticuloendotheliosis virus (from chicken line 151-5)				

ORGANISM	REFERENCE
cells of man, clone 713	
Reticuloendotheliosis virus	
viruses; Retroviridae: Mammalian type C	
retroviruses; 3-Reticuloendotheliosis virus group.	
1 (bases 510 to 602)	
Ridgway, A. A., Swift, R. A., Kung, H.-J. and Fujita, D. J.	
in vitro transcription analysis of the viral	

JOURNAL MEDLINE REFERENCE AUTHORS	TITLE
J. Virol. 54, 161-170 (1985) 85135063	B-lymphoma induction by reticuloendotheliosis virus: Characterization of a mutated chicken syncytial virus provirus involved in c-myc activation
2 (bases 1 to 887) Swift,R.A., Boerkoel,C.F., Ridgway,A., Fujita,D.J., Dodgson,J.B. and Kung,H.-J.	J. Virol. 61, 2084-2090 (1987) 87226389

FEATURES	
source	Location/Qualifiers 1. .887
LTR	/organism="Reticuloendotheliosis virus" /db_xref="taxon:11636" 246. .750 /note="3' LTR"
misc_feature	805. .806 /note="cryptic intron splice donor site"
BASE COUNT	212 a 216 c 227 g 232 t
ORIGIN	About 700 bp after segment 1.

Query Match	45.9%;	Score 431.8;	DB 81;	Length 887;
Best Local Similarity	66.8%;	Pred. No. 1.7e-126;		
Matches 360;	Conservative 137;	Mismatches 17;	Indels 25;	Gaps 3;

QY	2	augguggaaggagcucccgggggaauagcgucgugcguacucgcaucauuagcuucug	61
Db	245	ATGTGGGAGGGAGCTCCGGGGAAATAGCGCTGCCTACTACGCATATTAGCTTTC	304
QY	62	uaaucaugcuugcuugccuuaagccgcaccauugacuugaauauuuugcugauaucauuuc	121
Db	305	TAA TCATGCTTGCTTGCCCTTAGCCGCCCATTTGACTTGATATATTTCGCTGAT-----	356
QY	122	ucggaaucgycgaucauuucuggaaucgycgaucacaagaagcagcucucauagaccuuaaaaag	181
Db	357	-----ATCATTCTCGGAATCGGCATCAAGAAGCAGGCTCATTAACCATAAAGG	405
QY	182	aaauuguucguuggaagcgagcaucagaccacuugcgccaaccauacagagcaaacacga	241

D	b		406	AATGTTTGTGAAGGCAAGCATCAGACCCTTGACCCATCCATTCACGACAACACGA	465
O	y		242	gaucgaacuaucuaucugagccaaugguuaaaggcgagaugcuauccuccaaugaagg	301
D	b		466	GATCGAACTATCATCTAGAGGCCAATGGTTGTAAGGGCAGATGCTATCCTCCAATGAGGG	525
O	y		302	aaaaugucaugcaacaaccuguccuguuaagcggcguaauaaggccagugcaucucucgu	361
D	b		526	AAATGTCATGCACA-----TCCTGTAAGCGGCTATATAAGCCAGTCGATCTCTTGCT	580
O	y		362	cgggugcgccgucuaacacaauguuuugacgcgcgccagaucgaaucguuaaanaa	421
D	b		581	CGGGTCGCCGCTCTACACATTGTTGTGACGTGCGGCCCGATTCGATCTGTAATAAAA	640
O	y		422	guuuuuuuucuuaauaucuccuagauuuggcagugagagagaauuuuugucuguguaagc	481
D	b		641	G-CTTTTCTCTATATCCTCAGATTGGCAGTGAGAAGGAGATTGTTCGTGCTGTTGGC	699
O	y		482	uggcouacuugggugugguuaggggucgcgacugaaucgcguaguauuucgauacaacauu	540
D	b		700	TGGCCTACTGGGTGGGGTAGGGATCCGGACTGAATCCGTAGTATTTCGGTAATATATGT	758

RESULT	10	RESNVX	LOCUS	RESNVX	1530 bp	DNA	VRL	01-OCT-1996
DEFINITION				Spleen necrosis virus sequence from the end of the U3 region to 1.92kbp from the 5' end.				

ACCESSION	V01200 J02387
VERSION	V01200.1 GI:61757
KEYWORDS	integration site; provirus; terminal repeat.
SOURCE	Spleen necrosis virus.
ORGANISM	Spleen necrosis virus

REFERENCE AUTHORS TITLE JOURNAL MEDLINE	
	retroviruses; 1-Mammalian type C virus group. 1 (bases 2 to 470)
	Shimotohno, K., Mizutani, S. and Temin, H.M. Sequence of retrovirus provirus resembles that of bacter transposable elements
	Nature 285 (5766), 550-554 (1980)
	80254544

REFERENCE AUTHORS TITLE JOURNAL MEDICINE	2 (bases 1 to 1530) O'Rear, J.J. and Temin, H.M. Spontaneous changes in nucleotide sequence in proviruses of spleen necrosis virus, an avian retrovirus Proc. Natl. Acad. Sci. U.S.A. 79 (4), 1230-1234 (1982) 82174560
--	--

COMMENT
The meaning of the substitution at 739-749 is not completely clear in [2].

FEATURES	Location/Qualifiers
source	1. .1530
variation	/organism="Spleen necrosis virus" /db_xref="taxon:11836"
variation	313
variation	/note="T is missing in clone 13"
variation	314
variation	/note="G is missing in clone 13"
variation	315
variation	/note="G is missing in clone 13"
variation	327
variation	/note="T is missing in clone 13"
variation	367
variation	/note="G is A in clone 14-44"
variation	367
variation	/note="A may be G"
variation	374
variation	/note="C is T in clone 32"
variation	374
variation	/note="C is T in clone 13"
variation	375
variation	/note="T is C in clone 32"
variation	375
variation	/note="T is C in clone 13"

variation	383	/note="A is T in clone 32"
variation	383	/note="A is T in clone 13"
variation	474	/note="G is A in clone 13"
CDS	592..>1530	/note="reading frame (gag?)"
		/codon_start=1
		/protein_id="CAA24513.1"
		/db_xref="GI:61758"
		/db_xref="SWISS-PROT:P03342"
		/translation="MGOAGSGKGLTPLECILKNFSDEKKRAGDYGEDVDSFALRKLCE LEWPTFGVGPKEGTLDFKVVAAVRNLVEGNPGHPDQVIYITWTDITIERPKYLKSC GCKPHRTSKVLASOKVNPRLRPVLPASPSPPRIRRAQFLDERPLSPAPAPPPYPEV SAIVEDLYREGQDPDSTVMTSPPHTRSGLEFGAAGPSSGMPLRETGERDMTGRPMRTYV PFTSDILYNWKQNQNPSSSFSAQDPDQVLSLESEFYTHOPTWDDCQQLRLTLTTERER VRTESRREVRNDQGVQVTDEREIEAQPPATRPDWGS"
variation	594	/note="G is A in clone 63"
variation	732	/note="G is C in clone 13"
variation	739..749	/note="substitute with 135-188 in clone 13"
variation	949	/note="A is G in clone 63"
variation	1029	/note="C is A in clone 32"
variation	1085	/note="A is G in clone 13"
variation	1122	/note="A is G in clone 3-73"
variation	1192	/note="additional C in clone 3-73"
BASE COUNT	344 a	389 c 432 g 365 t
ORIGIN		

Query Match	45.3%;	Score 425.6;	DB 81;	Length 1530;
Best Local Similarity	64.5%;	Pred. No. 1.7e-124;		
Matches 387;	Conservative 160;	Mismatches 29;	Indels 24;	Gaps 8;

[illegible]

QY	821	agcgccocugcgaauuguguccacaccgcgcgycguugcgaaunaucuuugygagaguc	880
		: : : : : :	
Dd	477	AGCGCCTTGGCAATTGTTACCCACACC GCGGCTTGGAATAACTTTGGAGAGCC	536
QY	881	uuuugccuuccagugucuuuccguuuugacuquccucucuccucuccgcggccggyaugg	940
		::::: : : : : : : : : :	
Dd	537	TTTTGCTTCACAGTGTCTTCCGCTGTGTA CTGCTCCTCCTCCCTCCGCGCGGGGATGGG	596

RESULT	11
REXXX2	
LOCUS	REXXX2 859 bp DNA VRL 06-APR-1993
DEFINITION	Spleen necrosis virus (3' end) integrated in chicken cell.
ACCESSION	V01205 J02389
VERSION	V01205.1 GI:61794
KEYWORDS	
SOURCE	Spleen necrosis virus.
ORGANISM	Spleen necrosis virus

REFERENCE	
AUTHORS	1 (bases 1 to 859)
TITLE	Shimotohno, K., Mizutani, S. and Temin, H.M. Sequence of retrovirus provirus resembles that of bacterial transposable elements
JOURNAL	Nature 285 (5766), 550-554 (1980)

MEDLINE	00234344
FEATURES	Location/Qualifiers
source	1. .859

```

source      <1. 740
             /organism="Spleen necrosis virus"
             /proviral
             /db_xref="taxon:11836"
             741. .>859
source

```

BASE COUNT	245 a	199 c	214 g	201 t
ORIGIN				

Query Match	32.7%;	Score 307.6;	DB 81;	Length 859;
Best Local Similarity	62.8%;	Pred. No. 7.3e-87;		
Matches 323;	Conservative 112;	Mismatches 54;	Indels 25;	Gaps 7;

QY	31	gcugcucgcuuaacugccauuagcuucuguaaucaugcugcggaaggccag	90
Db	235	GCCCTGGCCACTAACCGGCATATTAGCTTCTGTACACATGCTTGCTGCCGTAGCCGCAT	294
QY	91	uguacuugaunauuuugcugauucauuuucucggaaucgcaucauuucucggaucg	150
Db	295	TGTACTTGAT-----ATGCCATTTCTCGAATCGCATCGCATCAAGTTTCG-----CT	338
QY	151	caucaagagcagcucuaagaccuauaaaaggaaauquuugyagcgagcaucagacc	210
Db	339	TCTCAGAGCAAGCCACAAACCACAAAAAGGAAACGCCACCGAAGGCAAGCATCAGACC	398
QY	211	acuuugcgccauccaaucaacgagcaaacaacgagaucgaacuaucauacu-gagccaauyu	269
Db	399	ACTTGCGCCATCCAATCATGTAACGGACACAGAGATCGACTATCATCTGAGCCCAATGGT	458
QY	270	uguuaaaggcgagaugcuaucuccuuccaaauaggygaaauugcauycacaucucgu--ccug	327
Db	459	TGTAAGAAGGCGAGATGCTACTCTCCCAATGAGGGAATAATGTATGTAACACCCCTGTAAAGCTG	518
QY	328	uaagcgcgcuaunaagccagguugcaucucuuugcucgggyucgccgucucuacaca---uug	384
Db	519	TAAAGCGCTATATAAGCCGGGTACATCTCTTGCTCGGGGTCCGCCCTGCACATTTGTTG	578
QY	385	uuugagcgcgcgcccagauucgaaucuguaanaaaauuuuuuucucucuaauccucag	444

LOCUS	DEFINITION	ACCESSION	RESULT 12
AF006066	1022 bp Fowlpox virus S envelope glycoprotein (env) 3' long terminal repeat, complete sequence.	AF006066	16-SEP-1997

SOURCE ORGANISM	Fowlpox virus. Fowlpox virus
-----------------	---------------------------------

viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Avipoxvirus.

REFERENCE	AUTHORS	TITLE	JOURNAL	VIROLOGY	235 (2), 367-376 (1997)
1 (bases 1 to 1022)	Hertig, C., Coupar, B. E., Gould, A. R. and Boyle, D. B.	Field and vaccine strains of fowlpox virus carry integrated sequences from the avian retrovirus, reticuloendotheliosis virus			

REFERENCE	2 (bases 1 to 1022)
AUTHORS	Hertig, C.H., Coupar, B.E.H., Gould, A.R. and Boyle, D.B.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAY-1997) Division of Animal Health, CSIRO, 5 Portarlington Road, Geelong, Victoria 3213, Australia
FEATURES	Location/Owner

/ SOURCE

gene

CDS

LTR

CDS

BASE COUNT	315 a	209 c	218 g	280 t
ORIGIN	ZNRN VDNFICFSNSIKQFIKELDLRSLSSSEHVIKLYGILDI"			

Query Match	17.18;	Score 160.4;	DB 82;	Length 1022;
Best Local Similarity	64.28;	Pred. No. 6.8e-40;		
Matches 140;	Conservative 53;	Mismatches 6;	Indels 19;	Gaps 1;

QY	1	aauuguggagggagcuccggygggaunagcgucgucgcuaacucgccaauuaagcuuu	60
Dd	429	AATGTGGAGGAGACTCCGGGGGAATAGCGCTGGCTCGCTAACTGCCATAATTAGCTTC	488
QY	61	guaaucaugcuugcuugccunagccgcaccauuguacuugaunaauuugcuganaucauu	120
Dd	489	GTAATCATGCTTGCTTGCCTTAGCCGCCAATTGTACTTGAATATTTTCGCTGAT-----	541
QY	121	cucggaaucgcaucauuucucggaucgycacaagaaggccuacauagaccanaaaag	180
Dd	542	-----ATCATTTCTCGGAATCGGCATCAAGACGAGCCTCATAAACCATAAAAG	589
QY	181	gaauugucguugaggcgagcaucaagaccacucgcgc	218
Dd	590	AAATGTTGTGAAGGCAAGCATCAGACCACCTTGCA	627

RESULT 13
AF006064

LOCUS	DEFINITION	2640 bp	DNA	VRL	16-SEP-1997
AF006064	Fowlpox virus M long terminal repeat complete				

kinase homolog gene, complete cds, and serpin gene, partial cds.
AF006064

AF006064.1 GI:2393888

SOURCE ORGANISM

viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Avipoxvirus.

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 2640)	Hertig, C., Coupar, B. E., Gould, A. R. and Boyle, D. B.	Field and vaccine strains of foot-and-mouth disease virus

sequences from the avian retrovirus, reticuloendotheliosis virus
 VIROLOGY 235 (2), 367-376 (1997)
 97428585

REFERENCE 2 (base

AUTHORS Herling, C.H., Coupar, B.E.H., Gould, A.R. and Boyle, D.B.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-1997) Division of Natural Sciences

JOURNAL

FEATURES	FEATURES
SOURCE	SOURCE
Location/Qualifiers	Location/Qualifiers
1..2640	1..2640

Source

CDS

```

LTR      905. .1152
          /note="remnant of LTR of reticuloendotheliosis virus"
CDS      1223. .2080

```

LTR

CDS

Tue Aug 8 08:33:17 2000

us-09-214-124-1.rge

Page 12

Query Match	12.5%;	Score 117.2;	DB 81;	Length 180;
Best Local Similarity	61.0%;	Pred. No. 3.7e-26;		
Matches	94;	Conservative	44;	Mismatches 13;
			Indels	3;
			Gaps	1;

[illegible]

Search completed: August 7, 2000, 10:53:33
Job time: 5987 sec

```
Run on:      August 7, 2000, 09:13:47 ; Search time 205.61 seconds
OM nuclelc - nuclelc search, using sw model (without alignments)
1143.819 Million cell updates/sec
```

US-09-214-124-1

Sequence	Title	perfect score
1 aaugggaggaggagcuccgg...	940	

Sequence IDENTITY-NUC Gapert 1.0
Scoring table: Gapop 10.0 , 1000 residues

Searched: 311585 seqs, 125096042
 Mapping chosen parameters: 623170

Total number of hits satisfactory

Minimum	DB seq	length:	0
Maximum	DB seq	length:	1000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : N_Geneseq_36:*
score. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

```

SUMMARIES

Result No.	Score	Match	length	DB	ID	Description
1	940	100.0	940	1	V63778	Sequence derived f
2	825.2	87.8	4643	1	X26113	Fowlpox virus (FPV)
3	578	61.5	578	1	V63779	Fowlpox derived f
4	437.8	46.6	3878	1	Q76041	Sequence derived f
5	436.2	46.4	512	1	X26109	Retrovirus vector
6	435.4	46.3	536	1	X26110	Chicken syncytial
7	430.4	45.8	5528	1	Q76038	Fowlpox virus (FPV)
8	420	44.7	5525	1	Q76039	Retrovirus vector
9	413	43.9	5519	1	Q76040	Retrovirus vector
10	160.4	17.1	296	1	X26111	Retrovirus virus (FPV)
11	160.4	17.1	1022	1	X26112	Fowlpox virus (FPV)
12	156	16.6	298	1	X26114	Fowlpox virus (FPV)
13	155	16.6	1100	1	X18925	Fowlpox virus (FPV)
14	33	3.5	1812	1	V74552	Phaseolus lunatus
15	32	3.4	5000	1	T62570	Staphylococcus aur
16	30.6	3.3	1276	1	T62572	Heat-resistant bar
17	30.6	3.3	2142	1	N60157	Heat-resistant bar
18	30.6	3.3	2205	1	N40081	Cattle growth horm
19	30.6	3.2	802	1	X20466	Sequence of genom
20	30.4	3.2	1382	1	Q64890	Human derived adre
21	29.8	3.2	11000	1	V21209-07	Continuation (8 of
22	29.8	3.1	32367	1	V35620	Human SHOX (short
23	29.2	3.1	13158	1	T75288	Nucleotide sequenc
24	28.8	3.1	3685	1	Q80523	Human mcl-1 gene.
25	28.8	3.1	3946	1	Q80523	Human sry-related
26	28.8	3.1	10266	1	T33007	Mouse mcl-1 gene.
27	28.8	3.1	11531	1	Q54222	BSSL/CEL Gene. DNA
28	28.6	3.0	4795	1	X32297	M. grisea PTH12 ge
29	28.6	3.0	11000	1	X20248-06	Continuation (7 of
30	28.6	3.0	111309	1	X20250	Borrelia burgdorfe
31	28.2	3.0	221	1	T65668	Polymorphic repeat
32	28.2	3.0	278	1	T65703	Polymorphic repeat
33						

45	27.8	3.0	1391	1	T79933
C 44	28	3.0	237326	1	V93927
C 43	28	3.0	235033	1	Q44362
C 42	28	3.0	9237	1	T47716
C 41	28	3.0	9237	1	Q67190
C 40	28	3.0	3412	1	V21209_14
C 39	28.2	3.0	1425	1	X13948
38	28.2	3.0	110000	1	Q51156
37	28.2	3.0	9636	1	X03046
36	28.2	3.0	9299	1	V57926
35	28.2	3.0	3789	1	V57903
34	28.2	3.0	3474	1	T07329
C 33	28.2	3.0	947	1	
C 32	28.2	3.0	947	1	
C 31	28.2	3.0	947	1	
C 30	28.2	3.0	947	1	
C 29	28.2	3.0	947	1	
C 28	28.2	3.0	947	1	
C 27	28.2	3.0	947	1	
C 26	28.2	3.0	947	1	
C 25	28.2	3.0	947	1	
C 24	28.2	3.0	947	1	
C 23	28.2	3.0	947	1	
C 22	28.2	3.0	947	1	
C 21	28.2	3.0	947	1	
C 20	28.2	3.0	947	1	
C 19	28.2	3.0	947	1	
C 18	28.2	3.0	947	1	
C 17	28.2	3.0	947	1	
C 16	28.2	3.0	947	1	
C 15	28.2	3.0	947	1	
C 14	28.2	3.0	947	1	
C 13	28.2	3.0	947	1	
C 12	28.2	3.0	947	1	
C 11	28.2	3.0	947	1	
C 10	28.2	3.0	947	1	
C 9	28.2	3.0	947	1	
C 8	28.2	3.0	947	1	
C 7	28.2	3.0	947	1	
C 6	28.2	3.0	947	1	
C 5	28.2	3.0	947	1	
C 4	28.2	3.0	947	1	
C 3	28.2	3.0	947	1	
C 2	28.2	3.0	947	1	
C 1	28.2	3.0	947	1	

ALIGNMENTS

uricase coding seq
 Novel human tumor
 Sequence of bovine
 Mouse bone morphog
 P. falciparum tran
 Continuation (15 o
 H. pylori GHPO 136
 Human P110 cDNA. R
 Human IL-1ra BAC C
 Human haemoch
 Hereditary haemoch
 Human calcium sign

10

ID	V63778 standard; DNA; 940 BP.
AC	V63778; 17-MAR-1999 (first entry) 5' end of REV-A; 5' end genomic RNA;
DT	Sequence derived from the type A; REV-A; 5' end genomic RNA;
DE	Reticuloendotheliosis virus type A; internal ribosome entry site; IRES; encapsidation;
KM	type C retrovirus; gene therapy; ss.
OS	retroviral vector; Reticuloendotheliosis virus.
PN	FR2762615-A1.
PD	30-OCT-1998. 005203.
PF	28-APR-1997; FR-005203.
PR	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA	Darlix Jr., Gabus Darlix C, Lopez LM;
PI	WP1; 99-037487/04.
DR	Expression vectors containing IRES and/or encapsidation enhancer
PT	derived from type C retrovirus other than FIMV and MOMLV
PS	Claim 5; Page 31-32; 43pp; French.
CC	The present sequence is derived from the 5' end of the genomic RNA of a retroviral
CC	Reticuloendotheliosis virus type A (REV-A). The specification also
CC	describes nucleotide sequence derived from all or part of the 5' end
CC	of the genomic RNA of a type C retrovirus other than Friend murine
CC	leukemia virus (FIMV) and Moloney murine leukemia virus (MOMLV). The
CC	5' derived sequences are used as an internal ribosome entry site (IRES)
CC	in a vector for permitting or enhancing the encapsidation of a retroviral
CC	vector. The vectors can be used for gene therapy, production of
CC	recombinant polypeptides or production of 260 G; 265 U;
SQ	Sequence 940 BP; 190 A; 225 C;

Query Match	100.0%; Score 940;	DB 1;	Length 940;	Gaps
Best Local Similarity	100.0%; Pred. No. 0;	Mismatches	Indels	
Matches 940; Conservative				
QY 1 aauguggagggaggcuccggygggaauagcgucgcguacacugccauuuagcuucu 120				
1 AAUGUGGAGGGAGGCUCGCCGGGGAUAGCGCUGCUCGUACACUGCCAUAUAGCUUCU 120				
DB 61 guaaucaugcungcuugccuaagccgcccauuguacuuaauuucgcugaaucauuu 180				
61 GUAAUCAUVGCTUGGCUUGCCUUAAGCCGCCAUUGUACUUAUAUUVUUUCGUGAUAUCAAUUU 180				
QY 121 cuuggaucggcaucauuucucggaucggcaucaagaagcaggcucuaaagaccuuaaag 240				
121 CUUGCAUUCGGCAUVCUUUCUGGAUUCGGCAUACAAGACGAGCUCUAUUGAACCAUAAAAG 240				
DB 181 gaauugucguugaggagcgagcaucaagaccacucugcgccaucaccauacgagacaacag 300				
181 GAUUGUCGUUGAGGAGCGAGCAUCAAGACCACUUGCGCCAUUCCAUCACGAGCAACACAG 300				
QY 241 agaucgaacuucauacucagagcccauuguuuaaaggsgaagucuaucuccaauagag 300				
241 AGAUCGAACUAUACUCAGAGCCCAUUGGUUUAAGGCGAGAUGCUAUUCCUCCAUGAGG 300				

us-09-214-124-1.rng

CC of FPV carry integrated REV sequences. Some of the FPV strains carry
CC near full length provirus of REV, and can give rise to infectious REV
CC when transfected into cell cultures and when chickens are infected. The
CC present invention provides the means by which a FPV vaccine free from
CC contamination by REV may be produced. Sequences X26110-115 represent
CC long terminal repeat sequences. The new recombinant vaccine free from
CC molecule can comprise a FPV genome. The new recombinant vaccine free from
SQ Sequence 16323

QY	7	6	5	4	3	2	1	0
Matches	638;	68.3%;	Score	825.2;	DB	1;	Length	4643;
Social Similarity	638;	68.3%;	Pred. No.	66-271;	Mismatch	253;		
Conservative	253;							

[illegible][illegible][illegible]

247 aaacuaucuaucugagccca
aaaccatttgcacatcattcacgagaca
aaagaagucg 246

307 qucauacaa-
306

[illegible][illegible]

1654 TTCTCTTCTATAATTTAAAG-CTT 1653

1717A

547 ucgucgggaauccucccccaucc-
:|||.|||||GATTCGTAATTTTCGGTACAACTATTTTCG
-55999C 546

607 guaaquaaacua...
606

[illegible][illegible]

CTGATTTGTTCCTTCAAGACGCYUGAUAAUUAUAACA
1953

[illegible]

accgcgcygcugcgaagaaagaaacuu...
|||||GTCAGACAGCGCCCTTGCAGATTGGTCTTACCGC
846

[illegible]

[illegible]

RESULT 5
X26109
ID X26109 standard; DNA; 512 BP.
AC X26109;
DT 20-MAY-1999 (first entry)
DE Chicken syncytial virus provirus REV LTR sequence.
KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
KM vaccine; chicken; LTR; long terminal repeat; ss.
OS Avian reticuloendotheliosis virus.
PN WO9907852-A1.
PD 18-FEB-1999.
PF 07-AUG-1998; AU0628.
PR 08-AUG-1997; AU-008454.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PI Boyle DB, Coupar BEH, Gould AR, Hertig C;
DR WPL; 99-167428/14.
PT Recombinant vaccine against fowlpox virus - is free of contamination
PT from reticuloendotheliosis virus, used to prevent the virus
PT occurrence in chickens
PT Example 4; Fig 5; 127pp; English.
PS The invention relates to a recombinant fowlpox virus (FPV) whose nucleic
CC acid does not encode an active reticuloendotheliosis virus (REV), and
CC whose genome does not contain any REV sequence. The recombinant FPV is
CC used to produce a vaccine against fowlpox virus. The vaccine is used to
CC prevent the occurrence of fowlpox virus in chickens. The chickens can be
CC vaccinated at older than 1 day of age. Alternatively, chickens older than
CC 1 day can be vaccinated with the vaccine subsequent to vaccination with
CC FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains
CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
CC near full length provirus of REV, and can give rise to infectious REV
CC when transfected into cell cultures and when chickens are infected. The
CC present invention provides the means by which a FPV vaccine free from
CC contamination by REV may be produced.
SQ Sequence 512 BP; 130 A; 113 C; 129 G; 140 T;

Query Match	46.48;	Score 436.2;	DB 1;	Length 512;
Best Local Similarity	67.48;	Pred. No. 7.6e-139;		
Matches 362;	Conservative 137;	Mismatches 13;	Indels 25;	Gaps 3;

QY	3	ugugggaggggagccuccggggggaauagcgcucgucgcguacacugcccauuuagcuucugu	62
Db	1	TGTGGAGGGGAGCTCCGGGGGAATAGCGCTGCCTTAACACTGCCATATTAGCTTCTGT	60
QY	63	aaucaugcuugcuugccuuaagccgccauuuaucuauauauuuucguganaucauuuu	122
Db	61	AATCATGCTTGCTTCCTTAGCCGCCCACTTGTTACTTGATATATTTCGTGAT-----	111
QY	123	cggaaucgycgaucuuuucucggaucgcgcgaucagaagcagcgcucacuagaaccacaaga	182
Db	112	-----ATCATTTCCTCGGAATCGGCATCAAGAGCAGCGCTCATTAACCATAAAGA	161
QY	183	aauguucguuuggagcgagcauccagaccacuuycgcaccauccaauacgagcaaacacgag	242
Db	162	AATGTTTGTGAAGGCAAGCATCAGACCACCTTGCAACCATCCAATCACGAACAACAACGAG	221
QY	243	aucgaacuaucuaucugagccaauugnuuuaaaggcagaugcuauccuccaauaggya	302
Db	222	ATCGAACTATCATACTGAGCCCAATGGTTGTAAAGGGCAGATGCTATCTCCAATGAGGGA	281

[illegible]

RESULT 6
X26110
ID X26110 standard; DNA; 536 BP.
AC X26110;
DT 20-MAY-1999 (first entry)
DE Fowlpox virus (FPV M5) LTR sequence.
KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
KM vaccine; chicken; LTR; long terminal repeat; ss.
OS Fowlpox virus.
PN W09907852-A1.
PD 18-FEB-1999.
PF 07-AUG-1998; AU0628.
PR 08-AUG-1997; AU-008454.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PI Boyle DB, Coupar BEH, Gould AR, Hertig C;
PT WPI; 99-167428/14.
PT Recombinant vaccine against fowlpox virus - is free of contamination
from reticuloendotheliosis virus, used to prevent the virus
occurrence in chickens
PS Claim 18; Fig 5; 127pp; English.
CC The invention relates to a recombinant fowlpox virus (FPV) whose nucleic
acid does not encode an active reticuloendotheliosis virus (REV), and
CC whose genome does not contain any REV sequence. The recombinant FPV is
CC used to produce a vaccine against fowlpox virus. The vaccine is used to
CC prevent the occurrence of fowlpox virus in chickens. The chickens can be
CC vaccinated at older than 1 day of age. Alternatively, chickens older than
CC 1 day can be vaccinated with the vaccine subsequent to vaccination with
CC FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains
CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
CC near full length provirus of REV, and can give rise to infectious REV
CC when transfectd into cell cultures and when chickens are infected. The
CC present invention provides the means by which a FPV vaccine free from
CC contamination by REV may be produced. Sequences X26110-115 represent FPV
CC long terminal repeat sequences. The new recombinant FPV nucleic acid
CC molecule can comprise a FPV LTR sequence selected from the above.
SQ Sequence 536 BP; 145 A; 115 C; 133 G; 143 T;

Query Match	46.3%;	Score 435.4;	DB 1;	Length 536;
Best Local Similarity	67.9%;	Pred. No. 1.5e-138;		
Matches 362;	Conservative 135;	Mismatches 11;	Indels 25;	Gaps 3;

[illegible]

Db	190	TTTGTGTGAGGCCAAGCATCAGACCACCTTGCACCATTCCAAATCAGGACAAACACGAGATCG	249
QY	247	aacuaucuaucugagcccaauguuuaaaggcgagaugcuauccccaugaggaaaaa	306
Db	250	AACATATCATACTGAGGCCAATGGTTGTAAGGGCAGATGCTATCCTCCAAATGAGGGAAAT	309
QY	307	gucaugcaacauccuguccuguaagcgcguaauaagaagccagugcaucucugcggg	366
Db	310	GTGATGCCAACA-----TCCTGTAAGCGGCTATATTAAGCCAGGTGCATCTCTTGCTCGGG	364
QY	367	ucgcggucuaacacauuguuugacgcgcgcgccagauucgaaucuguaaauaaaguuuu	426
Db	365	TGCGCGTCTACACATTTGTTGTGACGTCGCGGCCAGATTCGAATCTGTAATAAAG-CTT	423
QY	427	uuucuucauaauccucagauugcgagugaggaauuuuuuguuugugugugcgcgcc	486
Db	424	TTTCTTCTATATCCTCAGATTGGCAGTGAAGAGAGATTTTGTTCGTGCTGTGCGCTGGCC	483
QY	487	uacugggugggguaggguccggaucugaauccguauguauuucgaucacaacau	539
Db	484	TACTGGGTGGGGTAGGGATCCGGACTGAATCCGTAATTTCCGTACAACATT	536

RESULT	7
ID	076038 standard; DNA; 5528 BP.
AC	076038;
DT	20-JUL-1995 (first entry)
DE	Retrovirus vector pPOL11-R1.
KW	Retrovirus; vector; pPOL11-R1; spleen necrosis virus; SNV;
KW	cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW	long terminal repeat; LTR; encapsidation; gene transfer;
KW	gene therapy; ss.
OS	Spleen necrosis virus.
PN	WO9429437-A.
PD	22-DEC-1994.
PF	07-JUN-1994; U06415.
PR	07-JUN-1993; US-073345.
PA	(UYNE-) UNIV NEW JERSEY.
PI	Dornburg RC;
DR	WPI; 95-036467/05.
PT	Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT	ability to produce progeny virus, in helper cell which can infect
PT	host.cell and form provirus
PS	Example; page 13-14; 28pp; English.
CC	New recombination-free, highly efficient retroviral vectors
CC	pPOL11-R1 (given in Q76038), pPOL11-R2 (Q76039) and pPOL11-R3
CC	(Q76040) were obtained by replacing the U3 region of the left
CC	LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC	and extension of the encapsidation region. The vectors allow
CC	cell-type specific gene expression and eliminate risks of
CC	downstream activation of cellular proto-oncogenes.
SO	Sequence 5528 BP; 1265 A; 1423 C; 1457 G; 1383 T;

Query Match	45.8%;	Score 430.4;	DB 1;	Length 5528;
Best Local Similarity	62.4%;	Pred. No. 2.8e-136;		
Matches 397;	Conservative 169;	Mismatches 36;	Indels 34;	Gaps 8;

QY	333	ggcuaauaaagccaggugcaucucnugcugcgggugccgcguacaca-- -uugungug	389
Db	654	GTTCTATATAAGCAGAGCTCATCTTGTGCTCGGGGTCGCCGTCCACACATTGTTGTTG	713
QY	390	acgcgcggcccagaauucgaaucauaaaaaguuuuuuuucuauaucucagauug	449
Db	714	ACGTGC GGCC CAGATT CGAATCTGAATAAACTTTTTTTTCTGAATCC TCAGATTGG	773
QY	450	cagugaagaggaauuuugucgugugugagcuggccuacugguggg- guaggguccg	508
Db	774	CAGTGAGAGGAGATT TTTGTTCTGTTGCTGCTGCGCTACTGGGTGGCGCAGGATCCG	833
QY	509	gacugaaauccguaguuuucgaacaaauuuggggucucquccqda- uccucccca	567

Db	834	GACTGAATCCGTA	CTAGTACTTCGGTACACAACTTTGGGGGCTCGTCCGGGATACCCCTCCCAT	893
QY	568	cgcgcaagaugc	ccuacug-----uuucucgaacuccgycgcgguaaag-uaag	615
Db	894	CGGCAGAGGTG	CCCAACTGCTTCTTCGAACCTTCTTCGAATCCGGCGCGGTGAGTTAAG	953
QY	616	uacuuuauuuu	guuacucgcgagggguuugggaggaucgga---gugcggyacgucgcc	672
Db	954	TACTTGATTTT	TGGTACCTCGCGAGGCTTTGGGAGGATCGGAGTGGTGGCGGACGCTGCC	1013
QY	673	gggaagucca	ccaccucgcgcucagcaggggagcgcgccucgaucugagcuc-----ug	720
Db	1014	GGGAAGCTCC	ACACTCCGCTCAGACGGGACGCCCCTGACCTGACGCTCGAATTCAGATCTTG	1073
QY	721	ugguuacucgu	auuguuugugagaccgucuccaagacgguuauauaagucguuguuug	780
Db	1074	TGGTATCTGA	TTGTTGTTGAGCCGCTCCCTAAGACGGTGATA--CTAAGTCGTGGCTGTG	1131
QY	781	uguuuguuugu	uaaccuuguguuuguucgucacauugucgacagcgccucgcaauugugu	840
Db	1132	TGTTGTTGTT	TGCTTGTGTTGTTGTTGTTGTCGCTTGTGCGACAGCGCCTTGCGAATGCTGT	1191
QY	841	gcccaacacgc	gcgagcuugcgaaauaacuuugagagucuuuuugccuccacagucucucc	900
Db	1192	ACCACACACG	CGCGGCTTGGCAATAATACTTTGGAGAGCCTTTTGCCCTCCAGTGTCTCC	1251
QY	901	guuuguaucug	ucucucucuccucucgcgcggga	936
Db	1252	GTCGTACTCG	TCCCTCCTCTCCCTCCTCCGCGCGGGA	1287

```

RESULT      8
ID          Q76039
AC          Q76039;
DE          20-JUL-1995 (first entry)
DE          Retrovirus vector pPol11-R2.
KW          Retrovirus; vector; pPol11-R2; spleen necrosis virus; SNV;
KW          cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW          long terminal repeat; LTR; encapsidation; gene transfer;
KW          gene therapy; ss.
OS          Spleen necrosis virus.
PN          WO9429437-A.
PD          22-DEC-1994.
PE          07-JUN-1994; U06415.
PR          07-JUN-1993; US-073345.
PA          (UYNE-) UNIV NEW JERSEY.
PI          Dornburg RC;
DR          WPI; 95-036467/05.
PT          Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT          ability to produce progeny virus, in helper cell which can infect
PT          host cell and form provirus
PS          Example; Page 14-16; 28pp; English.
CC          New recombination-free, highly efficient retroviral vectors
CC          pPol11-R1 (given in Q76038), pPol11-R2 (Q76039) and pPol11-R3
CC          (Q76040) were obtained by replacing the U3 region of the left
CC          LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC          and extension of the encapsidation region. The vectors allow
CC          cell-type specific gene expression and elimination risks of
CC          downstream activation of cellular proto-oncogenes.
CC          Sequence 5525 BP; 1264 A; 1422 C; 1457 G; 1382 T.
SO          Sequence 5525 BP; 1264 A; 1422 C; 1457 G; 1382 T.

```

Query Match	44.7%	Score 420;	DB 1;	Length 5525;
Best Local Similarity	62.7%;	Pred. No. 9.9e-133;		
Matches 386;	Conservative 166;	Mismatches 30;	Indels 34;	Gaps 8;

[illegible]


```

      :||:||||| : : : : : ||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 731 TCTGTAATAAATTCTTTTCTGTAATCCTCAGATTGGCAGTGCAGAGAGATTGTT 790
QY 470 cguguguaagcgccuacucugugugg-guagggugccgagacugaaucguauuuc 528
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 791 CGTGGTGTGCTGGCCTACTGAGTGGGGCGCAGAGTCCGGAATCGGTAAGTACTTC 850
QY 529 gauacacauuuggggucugucggggu-uucccccacuggcagaaugccuacug-- 585
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 851 GGTACACACATTTGGGGCTCGTCCGGATACCTCCCATCCGAGAGGTGCCAACTGCT 910
QY 586 -----uuucucgaacuccgcgcgcguuag-uaaguacugauuuuguaucucg 635
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 911 TCTTGAACCTTCTTGAACCTCCGGCGCGGTGAGTTAAGTACTTGATTTTGTACTTCG 970
QY 636 cgaagguuugggagagucgga---guggcggagcugccggggaagcuccacucgcuc 692
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 971 CGAGGGTTGGGAGGATCGAGTGGTGGCGGAGCGCTGCCGGGAAGCTCCACCTCCGCTC 1030
QY 693 agcaggggagcgcgcgcugaucguc-----uguguaucugauuuugugug 740
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1031 AGCAGGGGAGCGCCCTGACCTGAGCTCGAATTCAGATCTTGATCTGATGTTGTTGA 1090
QY 741 accgucuccaagacgguauaauaauaagucuguguuuguuuguuuuguaucugug 800
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1091 GCCGTCCCTAAGACGGTGAFA--CTAAGTCGTGGCTGTGTGTTGTTGCTTGTG 1148
QY 801 uuuguuucgucacugucagacagcgcgcgcgaauguguguccacacgcgcgcucg 860
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1149 TTTGTTCTGCTGTTGTCAGACAGCGCCTTGCGAATTGGTGTACCCACACCGCGGCTTGC 1208
QY 861 gaauaauacuugagagagucuuuugccuccagugucucucguuuuguaucucucuc 920
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1209 GAATAATACCTTGGAGAGCCTTTTGCCTCCAGTGTCTTCCGTCTGTACTGCTCCTCT 1268
QY 921 cccucucgcgcgcgga 936
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1269 CCTCTCCGCGCGGA 1284

RESULT 9
076040
ID 076040 standard; DNA; 5519 BP.
AC 076040;
DT 20-JUL-1995 (first entry)
DE Retrovirus vector pPO111-R3.
KW Retrovirus; vector; pPO111-R3; spleen necrosis virus; SNV;
KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW long terminal repeat; LTR; encapsidation; gene transfer;
KW gene therapy; ss.
OS Spleen necrosis virus.
PN WO9429437-A.
PD 22-DEC-1994.
PF 07-JUN-1994; U06415.
PR 07-JUN-1993; US-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example; Page 16-17; 28pp; English.
CC New recombination-free, highly efficient retroviral vectors
CC pPO111-R1 (given in Q76038), pPO111-R2 (Q76039) and pPO111-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
SQ Sequence 5519 BP; 1264 A; 1420 C; 1456 G; 1379 T;

Query Match 43.9%; Score 413; DB 1; Length 5519;
```

```

      Best Local Similarity 62.9%; Pred. No. 2.4e-130;
Matches 383; Conservative 162; Mismatches 30; Indels 34; Gaps 8;

QY 360 cucggggugccgucucuaaca---uuguuugagcgcgcgcccagaauucgaucugua 416
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 672 CTCGGGGTCGCCCTCCGACATGTTGTTGACGTCGGGCCAGATTCCGAATCTGTAA 731
QY 417 uaaaauuuuuuucuaauaauccuacugauugcgagagaggaauuuugucgug 476
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 732 TAAACTTTTCTTCTGATCCCTCAGATTGGCAGTGAGAGAGAGATTGTTGCTGTG 791
QY 477 uaggcugccuacuguguggg-guagggugccgagcugaaucguauuuguaacaa 535
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 792 TTGCTGGCCTACTGGGTGGCGCAGGGATCCGGACTGAATCCGTACTCTCGGTACAA 851
QY 536 cauuuggggucugucggggu-uucccccacuggcagaaugcucuaucg----- 585
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 852 CATTTGGGGCTCGTCCGGGATACCTCCCATCCGAGAGGTGCCAACTGCTCTTCCA 911
QY 586 --uuucuucaacuccgcgcgcguuag-uaaguacugauuuuguaucucgcgaggu 642
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 912 ACTTCTTGAACCTCCGGCGCGGTGAGTTAAGTACTTGATTTTGGTACTCCGAGGGT 971
QY 643 uugggagaaucgga---guggcggagcgcgcggggaagcuccacucgcucagcagg 699
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 972 TTGGAGGATCGAGTGGTGGCGGAGCGCTGCCGGAAGCTCCACCTCCGCTCAGCAGGG 1031
QY 700 gacgcccugaucugagcuc-----uguguaucugauuuugugaacgcguc 747
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1032 GACGCCCTGACCTGACCTCGAATTCAGATCTTGTTGTTGTTGAGCCGCTCC 1091
QY 748 ccaagacgugauaauaauaagucuguguuuguuuguuuuguaucuuuguuuc 807
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1092 CTAAGACGGTGATA--CTAAGTCGTGGCTTGTGTGTTGTTGCTTGTGTTGTTG 1149
QY 808 gucauuugcagacagcgcgcgcgaauguguguccacacgcgcgcgcgcgcgaaau 867
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1150 GTCCGTTTGTGACAGAGCGCTTGCGAATTGTTGTACCCACACCGCGGCTTGCGAATAT 1209
QY 868 acuuugagagucuuuugccuccagugucucucguuuuguaucucucucucucuc 927
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1210 ACTTTGGAGAGCCTTTTGCTCCAGTGTCTCCGTCTGTACTGCTCCTCCTCCTCTC 1269
QY 928 cggccgcgga 936
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1270 CGGCGCGGA 1278

RESULT 10
X26111
ID X26111 standard; DNA; 296 BP.
AC X26111;
DT 20-MAY-1999 (first entry)
DE Fowlpox virus (FPV S3') LTR sequence.
KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
KW vaccine; chicken; LTR; long terminal repeat; ss.
OS Fowlpox virus.
PN WO9907852-A1.
PD 18-FEB-1999.
PF 07-AUG-1998; AU0628.
PR 08-AUG-1997; AU-008454.
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PI Boyle DB, Coupar BEH, Gould AR, Hertig C;
DR WPI; 99-167428/14.
PT Recombinant vaccine against fowlpox virus - is free of contamination
PT from reticuloendotheliosis virus, used to prevent the virus
PT occurrence in chickens
PS Claim 18; Fig 5; 127pp; English.
CC The invention relates to a recombinant fowlpox virus (FPV) whose nucleic
CC acid does not encode an active reticuloendotheliosis virus (REV), and
CC whose genome does not contain any REV sequence. The recombinant FPV is
CC used to produce a vaccine against fowlpox virus. The vaccine is used to
CC prevent the occurrence of fowlpox virus in chickens. The chickens can be
```


CC vaccinated at older than 1 day of age. Alternatively, chickens older than
CC 1 day can be vaccinated with the vaccine subsequent to vaccination with
CC FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains
CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
CC near full length provirus of REV, and can give rise to infectious REV
CC when transfectd into cell cultures and when chickens are infected. The
CC present invention provides the means by which a FPV vaccine free from
CC contamination by REV may be produced. Sequences X26110-115 represent FPV
CC long terminal repeat sequences. The new recombinant FPV nucleic acid
CC molecule can comprise a FPV LTR sequence selected from the above.
SO Sequence 296 BP; 75 A; 58 C; 81 G; 82 T;

```

Query Match          17.1%; Score 160.4; DB 1; Length 296;
Best Local Similarity 64.2%; Pred. No. 6.3e-45;
Matches 140; Conservative 53; Mismatches 6; Indels 19; Gaps 1;

QY      1 aauguggagggagcuccgsggggaauagcgucgucgcuaacucgcauuuagccucu 60
        ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       21 AATGTGGAGGAGCTCCGGGGGAATAGCGCTGGCTCGCTACTGCCATATTAGCTTCT 80

QY      61 guaaucaugcunugcunugccuuaggccgccauuguacuugaauauuuugcugauucauuu 120
        |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       81 GTATCATGCTTGCTGCCTTAGCCGCCATTGTACTTGATATAATTTCGCTGAT----- 133

QY      121 cucggaucgccaucuuuucucygaauucyggaucagcaucaagaagagcucuaagaccanaaag 180
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       134 -----ATCATTTTCTCGGAATCGGCATCAAGAGCAGGCTCATAAACCATAAAG 181

QY      181 gaaaugucguugagagggcgagcaucaagaccacucgvc 218
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       182 AAAATGTTGTGTAAGCCAAGCATCAGACCACCTTGCA 219

```

RESULT 11
 ID X26115 standard; DNA; 1022 BP.
 AC X26115;
 DT 20-MAY-1999 (first entry)
 DE Fowlpox virus (FPV S) 3' LTR sequence.
 KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
 OS vaccine; chicken; LTR; long terminal repeat; ss.
 PN Fowlpox virus.
 WO9907852-A1.
 PD 18-FEB-1999.
 PF 07-AUG-1998; AU0628.
 PR 08-AUG-1997; AU-008454.
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PI Boyle DB, Coupar BEH, Gould AR, Hertig C;
 WP1: 99-167428/14.
 PT Recombinant vaccine against fowlpox virus - is free of contamination
 from reticuloendotheliosis virus, used to prevent the virus
 occurrence in chickens
 PS Claim 18; Fig 6; 127pp; English.
 CC The invention relates to a recombinant fowlpox virus (FPV) whose nucleic
 acid does not encode an active reticuloendotheliosis virus (REV), and
 whose genome does not contain any REV sequence. The recombinant FPV is
 used to produce a vaccine against fowlpox virus. The vaccine is used to
 prevent the occurrence of fowlpox virus in chickens. The chickens can be
 vaccinated at older than 1 day of age. Alternatively, chickens older than
 1 day can be vaccinated with the vaccine subsequent to vaccination with
 FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains
 of FPV carry integrated REV sequences. Some of the FPV strains carry a
 near full length provirus of REV, and can give rise to infectious REV
 when transfected into cell cultures and when chickens are infected. The
 present invention provides the means by which a FPV vaccine free from
 contamination by REV may be produced. Sequences X26110-115 represent FPV
 long terminal repeat sequences. The new recombinant FPV nucleic acid
 molecule can comprise a FPV LTR sequence selected from the above.
 Sequence 1022 BP; 315 A; 209 C; 218 G; 280 T;

Query Match 17.18; Score 160.4; DB 1; Length 1022;

Best Local Similarity 64.2%; Pred. No. 1.3e-44;
Matches 140; Conservative 53; Mismatches 6; Indels 19; Gaps 1;

[illegible]

RESULT 12

ID X26112 standard; DNA; 298 BP.

AC X26112;

DT 20-MAY-1999 (first entry)

DE Fowlpox virus (FPV M3) LTR sequence.

KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;

KW vaccine; chicken; LTR; long terminal repeat; ss.

OS Fowlpox virus.

PN W09907852-A1.

PD 18-FEB-1999.

PF 07-AUG-1998; AU0628.

PR 08-AUG-1997; AU-008454.

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PI Boyle DB, Coupar BEH, Gould AR, Hertig C;

DR WPI; 99-167428/14.

PT Recombinant vaccine against fowlpox virus - is free of contamination

PT from reticuloendotheliosis virus, used to prevent the virus

PT occurrence in chickens

PS Claim 18; F19 5; 127P; English.

CC The invention relates to a recombinant fowlpox virus (FPV) whose nucleic

CC acid does not encode an active reticuloendotheliosis virus (REV), and

CC whose genome does not contain any REV sequence. The recombinant FPV is

CC used to produce a vaccine against fowlpox virus. The vaccine is used to

CC prevent the occurrence of fowlpox virus in chickens. The chickens can be

CC vaccinated at older than 1 day of age. Alternatively, chickens older than

CC 1 day can be vaccinated with the vaccine subsequent to vaccination with

CC FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains

CC of FPV carry integrated REV sequences. Some of the FPV strains carry a

CC near full length provirus of REV, and can give rise to infectious REV

CC when transfected into cell cultures and when chickens are infected. The

CC present invention provides the means by which a FPV vaccine free from

CC contamination by REV may be produced. Sequences X26110-115 represent FPV

CC long terminal repeat sequences. The new recombinant FPV nucleic acid

CC molecule can comprise a FPV LTR sequence selected from the above.

Sequence 298 BP; 84 A; 59 C; 73 G; 82 T;

Query Match	16.6%;	Score 156;	DB 1;	Length 298;
Best Local Similarity	64.6%;	Pred. NO. 2e-43;		
Matches 137; Conservative	51;	Mismatches	5;	Indels 19; Gaps 1.

```
QY      7   ggaggagcuccgagggaauagcguggcgcuaacugccauauuagcuucguaauc    66
        |||||:::|||||:::|||::|||::|||::|||::|||::|||::|||::|||
Db     29   GGAGGAGCTCCGGGGGAATAGCGCTGGCTGCCTAACTGCCATATTAGCTTGTATC    88
QY     67   auggcuugcugccuagccgccauuguacuugaauuuugcugaucauuuucucyga    126
        |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     89   ATGCTTGCTTGCCCTTAGCCGCACATTGTACTTGTATATTTCGCTGAT-----    135
QY    127   auggcaucauuucugcggaaucgycgaucacaagaagcaggcucauagaccuaaaaaggaaa    186
        |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    136   -----ATCATTTCTTCGGAATCGGCATCAAGACGAGGCTCATAAACCATAAAGAATG    189
```

987

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:13:46 ; Search time 172.61 Seconds
(without alignments)
749.102 Million cell updates/sec

Title: US-09-214-124-1
Perfect score: 940
Sequence: 1 aaugugggagggagcucg.....cccucuccggcggauggg 940

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	37.6	4.0	7218	1	US-08-232-463-14 Sequence 14, Appl
2	30.6	3.3	1276	4	US-08-793-599-1 Sequence 1, Appl1
3	30.6	3.3	2142	4	US-08-793-599-3 Sequence 3, Appl1
4	29	3.1	13158	4	US-08-687-080-105 Sequence 105, App
5	28.8	3.1	3934	5	US-09-226-568-18 Sequence 18, Appl
6	28.8	3.1	3946	1	US-08-077-848A-1 Sequence 1, Appl1
7	28.8	3.1	3946	5	US-09-211-640-1 Sequence 1, Appl1
8	28.8	3.1	3946	6	PCT-US94-03547-1 Sequence 1, Appl1
9	28.8	3.1	11531	1	US-08-068-945A-1 Sequence 1, Appl1
10	28.8	3.1	11531	2	US-08-442-806-1 Sequence 1, Appl1
11	28.2	3.0	221	1	US-08-222-177A-17 Sequence 17, Appl
12	28.2	3.0	278	1	US-08-222-177A-52 Sequence 52, Appl
13	28.2	3.0	9636	2	US-08-323-170B-1 Sequence 1, Appl1
14	28	3.0	3240	2	US-08-162-081B-34 Sequence 34, Appl
15	28	3.0	3240	3	US-08-780-872-34 Sequence 34, Appl
16	28	3.0	3412	2	US-08-162-081B-32 Sequence 32, Appl
17	28	3.0	3412	3	US-08-780-872-32 Sequence 32, Appl
18	28	3.0	246240	3	US-08-724-394A-20 Sequence 20, Appl
19	28	3.0	246240	3	US-08-724-394A-21 Sequence 21, Appl
20	28	3.0	246240	3	US-08-724-394A-22 Sequence 22, Appl
21	27.8	3.0	1391	1	US-08-261-662-1 Sequence 1, Appl1
22	27.8	3.0	1391	6	PCT-US95-07752-1 Sequence 1, Appl1
23	27.8	3.0	3566	5	US-08-689-421-32 Sequence 32, Appl
24	27.6	2.9	6545	4	US-08-843-530B-1 Sequence 1, Appl1
25	27.6	2.9	8224	4	US-09-010-398-14 Sequence 14, Appl
26	27.6	2.9	10803	5	US-09-080-044-1 Sequence 1, Appl1

27	27.4	2.9	943	5	US-08-705-875A-1	Sequence 1, Appl1
28	27.2	2.9	2055	1	US-07-842-349-1	Sequence 1, Appl1
29	27.2	2.9	2277	2	US-08-676-967-2	Sequence 2, Appl1
30	27.2	2.9	2277	2	US-08-676-974-2	Sequence 2, Appl1
31	27.2	2.9	2277	4	US-09-098-487-2	Sequence 11, Appl
32	27.2	2.9	2643	4	US-08-781-802-11	Sequence 11, Appl
33	27.2	2.9	3255	3	US-08-916-917-11	Sequence 11, Appl
34	27.2	2.9	3255	5	US-09-225-170-11	Sequence 9, Appl1
35	27.2	2.9	3545	4	US-08-781-802-9	Sequence 1, Appl1
36	27.2	2.9	3926	4	US-08-731-722-1	Sequence 2, Appl1
37	27.2	2.9	3926	4	US-08-731-722-2	Sequence 3, Appl1
38	27.2	2.9	3933	4	US-08-731-722-3	Sequence 1, Appl1
39	27.2	2.9	4315	4	US-08-781-802-1	Sequence 107, App
40	27.2	2.9	7208	5	US-09-166-186-107	Sequence 12, Appl
41	27	2.9	741	5	US-08-937-271-12	Sequence 16, Appl
42	27	2.9	1029	5	US-08-937-271-16	Sequence 1, Appl1
43	27	2.9	2281	2	US-08-164-614A-1	Sequence 1, Appl1
44	27	2.9	2281	4	US-08-456-489B-1	Sequence 1, Appl1
45	27	2.9	2281	6	PCT-US93-01720-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 13; Conservative 156; Mismatches 117; Indels 0; Gaps 0;	4.0%;	37.6;	DB 1;	7218;
QY 643	uuugaggaugcggaguggcggagcugccggaagcuccaccuccgcgucagcaggggac	702		
Db 1031	TTCCGAGCTTGCGCTGCAGGTCAGGAGGAGCTTGCGATYYYYYYYYYYYYYYYYY	1090		
QY 703	gccccugaucugagcucuguguaucugauuguuugugagccgucuccaagacgugaua	762		
Db 1091	YYY	1150		
QY 763	uauaagucguguguuuguguguuuuguaaccuuuguuugucacuuugcagacag	822		
Db 1151	YYY	1210		
QY 823	cgccccugcgaaugugugugccacacccgcgcgcuugcgaaauaauacuuuggagagucuu	882		
Db 1211	YYY	1270		
QY 883	uugccuccagugucucccguaucucgucuccuccuccucc	928		
Db 1271	YYY	1316		

```

RESULT 2
US-08-793-599-1/c
; Sequence 1, Application US/08793599
; Patent No. 5952489
; GENERAL INFORMATION:
; APPLICANT: OKADA, YUKIO
; APPLICANT: YOSHIGI, NAOHIRO
; APPLICANT: ITO, KAZUTOSHI
; APPLICANT: KIHARA, MAKOTO
; TITLE OF INVENTION: TISSUE-SPECIFIC PROMOTER
; FILE REFERENCE: 2589-0056-0 PCT
; CURRENT APPLICATION NUMBER: US/08/793, 599
; CURRENT FILING DATE: 1997-03-04
; EARLIER APPLICATION NUMBER: PCT/JP96/01866
; EARLIER FILING DATE: 1996-07-05
; EARLIER APPLICATION NUMBER: JP 191028/1995
; EARLIER FILING DATE: 1995-07-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Barley
US-08-793-599-1

```

[illegible]

RESULT 3
US-08-793-599-3/c
; Sequence 3, Application US/08793599
; Patent No. 5952489
; GENERAL INFORMATION:
; APPLICANT: OKADA, YUKIO
; APPLICANT: YOSHIGI, NAOHITO
; APPLICANT: ITO, KAZUTOSHI

```

; APPLICANT: KIHARA, MAKOTO
; TITLE OF INVENTION: TISSUE-SPECIFIC PROMOTER
; FILE REFERENCE: 2589-0056-0 PCT
; CURRENT APPLICATION NUMBER: US/08/793, 599
; CURRENT FILING DATE: 1997-03-04
; EARLIER APPLICATION NUMBER: PCT/JP96/01866
; EARLIER FILING DATE: 1996-07-05
; EARLIER APPLICATION NUMBER: JP 191028/1995
; EARLIER FILING DATE: 1995-07-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Barley
US-08-793-599-3

```

[illegible]

```

RESULT      4
; US-08-687-080-105/c
; Sequence 105, Application US/08687080
; Patent No. 5965427
;
; GENERAL INFORMATION:
;   APPLICANT: Gregory Dolganov
;   TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
;   NUMBER OF SEQUENCES: 175
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Dehlinger & Associates
;     STREET: 350 Cambridge Avenue, Suite 250
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94306
;
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/687,080
;   FILING DATE: 17-JUL-1996
;   CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/592,126
;   FILING DATE: 26-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sholtz, Charles K.
;     REGISTRATION NUMBER: 38,615
;     REFERENCE/DOCKET NUMBER: 4600-0111.30
;
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 324-0880
;   TELEFAX: (415) 324-0960
;
; INFORMATION FOR SEQ ID NO: 105:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 13158 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;

```



```

;      ANTI-SENSE:  NO
;      ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  5' END OF INTRON 21 OF RAD50 GENOMIC
;      INDIVIDUAL ISOLATE:  SEQUENCE
;
US-08-687-080-105

```

Query Match	3.18;	Score 29;	DB 4;	Length 13158;
Best Local Similarity	34.08;	Pred. No. 20;		
Matches 18;	Conservative 20;	Mismatches 15;	Indels 0;	Gaps 0;

[illegible]

RESULT 5
US-09-226-568-18/c
; Sequence 18, Application US/09226568
; Patent No. 6001992

```

; CURRENT INVENTION:
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcusson, Eric G.
; TITLE OF INVENTION: Antisense Modulation of No. 6001992e1 Anti-apoptotic
; TITLE OF INVENTION: bcl-2-Related Proteins
; FILE REFERENCE: ISPH-0337
; CURRENT APPLICATION NUMBER: US/09/226,568
; CURRENT FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3934
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
; US-09-226-568-18

```

Query Match	3.18;	Score 28.8;	DB 5;	Length 3934;
Best Local Similarity	56.28;	Pred. No. 11;		
Matches 27; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0

Oy 887 cuccagugucuuaccguuuuaguacuguccuucccuuccgagcg 934
| : | | | | : | : | : | : | : | : | : | : | : | : |
Db 600 CTCACGCACTGCCCGGTACAACTCCTTCCTCCTCCTCTGCTGGCG 553

```

RESULT      6
US-08-077-848A-1/c
: Sequence 1, Application US/08077848A
: Patent No. 5470955
:
: GENERAL INFORMATION:
:
: APPLICANT: Craig, Ruth W.
:
: TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
:
: TITLE OF INVENTION: POLYPEPTIDE
:
: NUMBER OF SEQUENCES: 4
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
:
: STREET: 1880 Century Park East, Suite 500
:
: CITY: Los Angeles
:
: STATE: California
:
: COUNTRY: USA
:
: ZIP: 90067
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:

```

```

1  APPLICATION NUMBER: US/08/077,848A
2  FILING DATE: 16-JUN-1993
3  CLASSIFICATION: 424
4  ATTORNEY/AGENT INFORMATION:
5  NAME: Hallé, Ph.D., Lisa A.
6  REGISTRATION NUMBER: 38,347
7  REFERENCE/DOCKET NUMBER: PD-2845
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE: (619) 455-5100
10 TELEFAX: (619) 455-5110
11 INFORMATION FOR SEQ ID NO: 1:
12 SEQUENCE CHARACTERISTICS:
13     LENGTH: 3946 base pairs
14     TYPE: nucleic acid
15     STRANDEDNESS: single
16     TOPOLOGY: linear
17     MOLECULE TYPE: DNA (genomic)
18     IMMEDIATE SOURCE:
19     CLONE: mcl-1
20     FEATURE:
21     NAME/KEY: CDS
22     LOCATION: 61..1110
23     OTHER INFORMATION: /note= "When nucleotide 740 = C,
24     OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
25     OTHER INFORMATION: acid 227 = V."
26
27 US-08-077-848A-1

```

Query Match	3.1%	Score 28.8;	DB 1;	Length 3946;
Best Local Similarity	56.2%;	Pred. No. 11;		
Matches 27; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0.

Dy 887 cuccagugucuccgubuuuacucgcucuccucucccucucccgccgg 934
|:|:| | : |:| : |:|:|:|:|:| :|:|:| :| | |
Db 600 CTCACGACTGCCGTACAACCTCGTCTCCTCTCTGCTGGCGG 553

```

RESULT 7
US-09-211-640-1/c
; Sequence 1, Application US/09211640
; Patent No. 6020466
;
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,640
; FILING DATE:
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/441,375
; FILING DATE:
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
;
; INFORMATION FOR SEQ ID NO: 1:

```



```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note="When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
US-09-211-640-1

Query Match          3.1%; Score 28.8; DB 5; Length 3946;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 887 cuccagugucuuuacugucucucucucucucucgcccgcgcg 934
      ||||| | | | | | | | | | | | | | | | | | | | |
Db 600 CTCACGCGACTGCCGGTACACTCTCTCTCTCTCTCTGTCGCGG 553

RESULT 8
PCT-US94-03547-1/c
; Sequence 1, Application PC/TUS9403547
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: MYELOID CELL LEUKEMIA ASSOCIATED GENE
; TITLE OF INVENTION: MCL-1
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03547
; FILING DATE: 31-MAR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note="When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
PCT-US94-03547-1
```

```
Query Match          3.1%; Score 28.8; DB 6; Length 3946;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 887 cuccagugucuuuacugucucucucucucucucgcccgcgcg 934
      ||||| | | | | | | | | | | | | | | | | | | | |
Db 600 CTCACGCGACTGCCGGTACACTCTCTCTCTCTCTCTGTCGCGG 553

RESULT 9
US-08-068-945A-1
; Sequence 1, Application US/08068945A
; Patent No. 5616483
; GENERAL INFORMATION:
; APPLICANT: Bjursell, Gunnar
; APPLICANT: Carlsson, Peter
; APPLICANT: Enerback, Sven
; APPLICANT: Hansson, Lennart
; APPLICANT: Lidberg, Ulf
; APPLICANT: Nilsson, Jeanette
; APPLICANT: Tornell, Jan
; TITLE OF INVENTION: New DNA Sequences
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,945A
; FILING DATE: 27-MAY-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9201809-2
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9201826-6
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9202088-2
; FILING DATE: 03-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9300902-5
; FILING DATE: 19-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)819-8783
; TELEFAX: (212)354-8113
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Mammary gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
```



```

LOCATION:      ..4904, 6193..6323, 6501..6608, 6751..6868, 8333
LOCATION:      ..8521, 8719..8922, 10124..10321, 10650..11394)
FEATURE:
NAME/KEY:    mat_peptide
LOCATION:      join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION:      ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION:      ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION:
/EC_number= 3.1.1.1
/product= "Bile Salt-stimulated Lipase"
FEATURE:
NAME/KEY:    5'UTR
LOCATION:      1..1640
FEATURE:
NAME/KEY:    TATA_signal
LOCATION:      1611..1617
FEATURE:
NAME/KEY:    exon
LOCATION:      1641..1727
FEATURE:
NAME/KEY:    exon
LOCATION:      4071..4221
FEATURE:
NAME/KEY:    exon
LOCATION:      4307..4429
FEATURE:
NAME/KEY:    exon
LOCATION:      4707..4904
FEATURE:
NAME/KEY:    exon
LOCATION:      6193..6323
FEATURE:
NAME/KEY:    exon
LOCATION:      6501..6608
FEATURE:
NAME/KEY:    exon
LOCATION:      6751..6868
FEATURE:
NAME/KEY:    exon
LOCATION:      8335..8521
FEATURE:
NAME/KEY:    exon
LOCATION:      8719..8922
FEATURE:
NAME/KEY:    exon
LOCATION:      10124..10321
FEATURE:
NAME/KEY:    exon
LOCATION:      10650..11490
FEATURE:
NAME/KEY:    3'UTR
LOCATION:      11491..11531
;
US-08-068-945A-1

```

Query Match	3.18;	Score 28.8;	DB 1;	Length 11531;
Best Local Similarity	33.68;	Pred. No. 21;		
Matches 51; Conservative	24;	Mismatches 77;	Indels 0;	Caps 0;

QY	667	gcugccgaggagcuccaccuccgcgucacgacgagggagcgcuccugaucugagcucucuguan	726
Db	7107	gctcactgcacaacctccaccctccctgggttcaagtgattctctgactcaacctcccatgtag	7166
QY	727	cugauuguuugugagaccgucuccaagacgugauaauaagucuguguuuguguuug	786
Db	7167	ctgggactacagggcacacatgccaccactgccacagataattttctctggttagtaggcat	7226
QY	787	uuuguaaccuuguguuugucgucacuuugc	818
Db	7227	ggagtttcattcgtgttagctaggatgatctcg	7258
RESULT	10		
US-08-442-806-1			

```

Sequence 1, Application US/08442806
Patent No. 5716817

GENERAL INFORMATION:
APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Enerback, Sven
APPLICANT: Hansson, Lennart
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
APPLICANT: Tornell, Jan
TITLE OF INVENTION: Genomic DNA Sequences
TITLE OF INVENTION: Encoding Human BSSL/CEL
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,806
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,945
FILING DATE: 27-MAY-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9202088-2
FILING DATE: 03-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9300902-5
FILING DATE: 19-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)819-8783
TELEFAX: (212)354-8113
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
OTHER INFORMATION: /EC_number= 3.1.1.1

```



```

? OTHER INFORMATION: /product- "Bile Salt-Stimulated Lipase"
? FEATURE:
? NAME/KEY: 5'UTR
? LOCATION: 1..1640
? FEATURE:
? NAME/KEY: TATA_signal
? LOCATION: 1611..1617
? FEATURE:
? NAME/KEY: exon
? LOCATION: 1641..1727
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4071..4221
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4307..4429
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4707..4904
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6193..6323
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6501..6608
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6751..6868
? FEATURE:
? NAME/KEY: exon
? LOCATION: 8335..8521
? FEATURE:
? NAME/KEY: exon
? LOCATION: 8719..8922
? FEATURE:
? NAME/KEY: exon
? LOCATION: 10124..10321
? FEATURE:
? NAME/KEY: exon
? LOCATION: 10650..11490
? FEATURE:
? NAME/KEY: 3'UTR
? LOCATION: 11491..11531
?
US-08-442-806-1

```

Query Match	3.1%;	Score 28.8;	DB 2;	Length 11531;
Best Local Similarity	33.6%;	Pred. No. 21;		
Matches	51;	Conservative	24;	Mismatches 77; Indels 0; Gaps 0;
QY 667	gcugccgggaagcuccaccucgcgcucagcaggggagcggcccgauaucgacucucuguguan	726		
Db 7107	GCTCACTGCAACCTCCACCCTCTGGGTTCAAGTAGATTCTCTGACTCAACCTCCCATGTAG	7166		
QY 727	cugauuuguuugugaccgcgucuccaagacgsgugauaauaauaagucyuguuuguguguuug	786		
Db 7167	CTGGGACTACAGGCACACATGCCACCACCATGCCACAGATAATTCTCTGTGTGTTAGTAGGAT	7226		
QY 787	uuuguaaccuuguguuuguuugucacauugucg	818		
Db 7227	GGAGTTTCATCGTGTAGCTAGGATGATCTCG	7258		

RESULT 11
 US-08-222-177A-17
 : Sequence 17, Application US/08222177A
 : Patent No. 5582979
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Weber, James L.
 : TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
 : TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
 : NUMBER OF SEQUENCES: 460
 :
 : CORRESPONDENCE ADDRESS:
 :

```

ADDRESSSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
INDIVIDUAL ISOLATE: Caucasian
TISSUE TYPE: Blood
IMMEDIATE SOURCE:
CLONE: Mfd30
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 3
FEATURE:
NAME/KEY: repeat_region
LOCATION: 103..139
OTHER INFORMATION: /rpt_type="tandem"
OTHER INFORMATION: /rpt_family="(dc-da)n.(dc-dt)n"
OTHER INFORMATION: /citation="(12)"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 72..91
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /standard_name="PCR primer"
OTHER INFORMATION: /citation="(11)"
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (144..163)
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /standard_name="PCR primer"
OTHER INFORMATION: /citation="(11)"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..221
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /standard_name="Only one strand sequenced"
PUBLICATION INFORMATION:
AUTHORS: Weber, J. L.
AUTHORS: May, P. E.
TITLE: Dinucleotide repeat polymorphism at the
TITLE: D3S240 locus
JOURNAL: Nucleic Acids Res.

```


VOLUME: 18
PAGES: 2201-
DATE: 1990
PUBLICATION INFORMATION:
AUTHORS: Weber, James L.
TITLE: Abundant Class of Human DNA Polymorphisms
TITLE: Which Can Be Typed Using the Polymerase Chain
TITLE: Reaction
JOURNAL: Am. J. Hum. Genet.
VOLUME: 44
PAGES: 388-396
DATE: 1989
US-08-222-177A-17

Query Match	3.0%;	Score 28.2;	DB 1;	Length 221;
Best Local Similarity	28.9%;	Pred. No. 3.3;		
Matches	28;	Conservative	26;	Mismatches 43;
				Indels 0;
				Gaps 0;

OY 725 aucugauuguuugagccgucuccaaagacgcgugauaauaagaucgugguuuguguguu 784
| : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 55 ATGGAATGTGGTGTTCGAATCAGTGATGCATGCACATATATACTGTGTGTGTG 114

Dy 785 uguauguaacaauguaugucagcaaaugucgaca 821
::: ::::: : | : | |
Db 115 TGTGTGTGTGTGTGTGTGTGTTAAATGTAGACA 151

RESULT 12
US-08-222-177A-52

GENERAL INFORMATION:
 APPLICANT: Weber, James L.
 TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
 TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
 NUMBER OF SEQUENCES: 460
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dewitt Ross & Stevens, S.C.
 STREET: 8000 Excelsior Drive, Suite 401

;; CITY: Madison
;; STATE: Wisconsin
;; COUNTRY: USA

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

FILING DATE: 435
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/341,562
 FILING DATE: 21-APR-1989
 ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:

```

; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
;

```

```

1 ORGANISM: Homo sapiens
2 INDIVIDUAL ISOLATE: Caucasian
3 TISSUE TYPE: Blood
4 IMMEDIATE SOURCE:
5 CLONE: M6154
6 POSITION IN GENOME:
7 CHROMOSOME/SEGMENT: 5q
8 FEATURE:
9 NAME/KEY: repeat_region
10 LOCATION: 76..118
11 OTHER INFORMATION: /rpt_type="tandem"
12 OTHER INFORMATION: /rpt_family="(dc-da)n.(dg-dt)n"
13 OTHER INFORMATION: /citation="(12)"
14 FEATURE:
15
16

```

```

; NAME/KEY: misc-feature
; LOCATION: 20..39
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence- EXPERIMENTAL
; OTHER INFORMATION: /standard_name= "PCR primer"
; OTHER INFORMATION: /citation- ([1])

```

```
;
;
FEATURE:
    NAME/KEY:      misc.feature
    LOCATION:       complement (198..217)
    IDENTIFICATION METHOD: experimental
    OTHER INFORMATION: /evidence=EXPERIMENTAL
    ;
    OTHER INFORMATION: /standard_name="PCR primer"
    ;
    OTHER INFORMATION: /citation=(11)
```

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..278
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence- EXPERIMENTAL
; OTHER INFORMATION: /standard_name= "Only one strand sequenced"
; PUBLICATION INFORMATION:
;
```

;
AUTHORS: Weber, J. L.
AUTHORS: Polymeropoulos, M. H
AUTHORS: May, P. E.
AUTHORS: Kwiktek, A. E.
AUTHORS: Xiao, H.
AUTHORS: McPherson, J. D.
;

AUTHORS: Wasmuth, J. J.
TITLE: Mapping of human chromosome 5 microsatellite
TITLE: polymorphisms
JOURNAL: Genomics
DATE: 1991

DATE: 1991
PUBLICATION INFORMATION:
AUTHORS: Weber, James L.
AUTHORS: May, Paula E.
TITLE: Abundant Class of Human DNA Polymorphisms
TITLE: Which Can Be Typed Using the Polymerase Chain
TITLE: Reaction
JOURNAL: Am J Hum Genet

US-08-222-177A-52

Query Match	3.0%;	Score 28.2;	DB 1;	Length 278;
Best Local Similarity	30.5%;	Pred. No. 3.8;		
Matches 32; Conservative	25;	Mismatches 48;	Indels 0;	Gaps 0;

Qy 705 ccugaucugagcucucuguguaucugauuguuguggaccgcucuccaagacgugauaau 76
|| :|:| | | :|:| : | | | | :|
Db 4 CCATTTCTGGGCGCACATGTATCTAGCCATGTGTACACAGGCCGGGAAGCTTGTGCTGG 63

[illegible]

RESULT 13
US-08-323-170B-1/c


```

: Sequence 1, Application US/08323170B
: Patent No. 5733772
: GENERAL INFORMATION:
: APPLICANT: Williamson, Kim C.
: APPLICANT: Kaslow, David C.
: TITLE OF INVENTION: Cloning and Expression of Plasmodium
: TITLE OF INVENTION: falciiparum Transmission-Blocking Target Antigen, Pfs230
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/323,170B
: FILING DATE: 13-OCT-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/010,409
: FILING DATE: 29-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Quine, Jonathan A.
: REGISTRATION NUMBER: P-41,261
: REFERENCE/DOCKET NUMBER: 015280-113100US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 149..9556
: US-08-323-170B-1

```

```

Query Match          3.0%; Score 28.2; DB 2; Length 9636;
Best Local Similarity 31.5%; Pred. No. 31;
Matches 28; Conservative . 23; Mismatches 38; Indels 0; Gaps 0;

QY 396 ggcccagaauucgaaucauuaaaaguuuuuuuucucuaauccuagauugcgaguga 455
      ||||| : : : ||| : : : : : : : ||| : ||| : |
Db 4797 GGTCTATATGTTGATTTGACATTATATTTTTTTTTTTAAATCAACACATATACATTTA 4738
      ||||| : : : : : : : : : : : : : : : : :

QY 456 gagagauuuuugucguugugugagcgug 484
      ||||| : : : ||| : ||| : ||
Db 4737 AATGATATATCATTTATTCATTAAAGCTGG 4709
      ||||| : : : ||| : ||| : ||

RESULT 14
US-08-162-081B-34
; Sequence 34, Application US/08162081B
; Patent No. 5824492
;
GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:

```

```

1 ADDRESSSEE: Felte & Lynch
2 STREET: 805 Third Avenue
3 CITY: New York
4 STATE: New York
5 COUNTRY: USA
6 ZIP: 10022
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
10
11 COMPUTER: IBM PS/2
12 OPERATING SYSTEM: PC-DOS
13 SOFTWARE: Wordperfect
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/162, 081B
17 FILING DATE: February 7, 1994
18
19 CLASSIFICATION: 435
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: PCT/GB93/00761
23
24 FILING DATE: 13 April 1993
25
26 ATTORNEY/AGENT INFORMATION:
27
28 NAME: Pasqualini, Patricia A.
29 REGISTRATION NUMBER: 34,894
30 REFERENCE/DOCKET NUMBER: LUD 5256
31
32 TELECOMMUNICATION INFORMATION:
33
34 TELEPHONE: (212) 688-9200
35
36 TELEFAX: (212) 838-3884
37
38 INFORMATION FOR SEQ ID NO: 34:
39
40 SEQUENCE CHARACTERISTICS:
41
42 LENGTH: 3240 base pairs
43 TYPE: nucleic acid
44 STRANDEDNESS: single
45 TOPOLOGY: linear
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
```

	Query Match	Best Local Similarity	Score	DB 2;	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	766	aagucgugguuuuguguuuuacccuuguguuugucacauugcagacgcgc	825							
Db	2475	AAATCAAGGCTTGTGATCTTCGAATGTTACCTTATGGTTGTCTCATAATCGGTGACTGTGT	2534							
QY	826	ccugcgaaauugugug	841							
Db	2535	GGGACTTATTGAGGTG	2550							

RESULT 15
US-08-780-872-34
; Sequence 34, Application US/08780872
; Patent No. 5846824
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayiotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780, 872

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:13:46 ; Search time 3871.63 Seconds
(without alignments)
1070.670 Million cell updates/sec

Title: US-09-214-124-1
Perfect score: 940
Sequence: 1 aauguggagggagcuccg.....ccucuccgcccgggauggg 940

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: em_est35: *
83: em_est36: *
84: em_est37: *
85: gb_est48: *
86: gb_est49: *
87: gb_est50: *
88: gb_est51: *
89: gb_est52: *
90: gb_est53: *
91: gb_est54: *
92: gb_est55: *
93: gb_gss1: *
94: gb_gss2: *
95: gb_gss3: *
96: gb_gss4: *
97: em_gss1: *
98: em_gss2: *
99: em_gss3: *
100: em_gss4: *
101: gb_gss5: *
102: gb_gss6: *
103: gb_gss7: *
104: gb_gss8: *
105: gb_gss9: *
106: em_gss5: *
107: em_gss6: *
108: em_gss7: *
109: em_gss8: *
110: em_gss9: *
111: em_gss10: *
112: em_gss11: *
113: gb_gss10: *
114: gb_gss11: *
115: em_gss12: *
116: gb_gss12: *

[illegible][illegible]

fly), genomic survey sequence.
ACCESSION AL106358
VERSION AL106358.1 GI:5621512
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1225)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.
FEATURES
source 1..1225
/organism="Drosophila melanogaster"
/plasmid="pBelOBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN15C20"
/note="end : SP6"
BASE COUNT 84 a 138 c 318 g 290 t 395 others
ORIGIN
Query Match 4.3%; Score 40.2; DB 123; Length 1225;
Best Local Similarity 15.4%; Pred.No.0.17;
Matches 76; Conservative 171; Mismatches 243; Indels 5; Gaps 1;
QY 347 ggugcaucucucgucgugcgucgcgucacacauugugagcgcgccacgaauuc 406
Db 651 KGSSTTTTGGSGGGKKTTGGCCYTTTCTTKTKTCTBBBBSBBBBSBSCSS 710
QY 407 gaaucuguaaauuu 466
Db 711 SKBYTYCGGKGRWTTTWTMTGTTKTKRMTTTTTTTKTKGGGGGGGDTGWT 770
QY 467 guucgugugugagcgucgucacucguguggggaggggucgagcagaaucguauu 526
Db 771 GTTGGKGGTTTTTTTGTGTTTCTTKTKGTTGGGKGKGGGGGGGGGGGGGGGG 830
QY 527 ucgaucacaauuuggggucgucggaauucccaccagcgagaagucgucacugu 586
Db 831 GTKGT----KTKKGTGTGTGTTGTTGTTGKKKKGTKKGGKGGKGGKGGKGGT 885
QY 587 uucucgaacucgcgcgcggaagaaagaaucuuuuuuuuuuuuuuuuuuuuuuuu 646
Db 886 TGGKGGKGG 945
QY 647 gaggaucgagagugcgagcgagcgaggaucacacucgagcagcgagggagcccc 706
Db 946 GTKGG 1005
QY 707 ugaucugagcucuguguaucugauugugagcagcucuccaagaagcgugauaaua 766
Db 1006 TKTITTKWKGTTGTTTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1065
QY 767 agucguguguguguguuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuu 826
Db 1066 KKGCGTTKGTGKTKGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1125
QY 827 cuagcaauugugug 841

Db 1126 GTGGGGGTGTGTGKK 1140
RESULT 5
CNS017GS
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN17J04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL108022
VERSION AL108022.1 GI:5628326
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.
FEATURES
source 1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelOBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN17J04"
/note="end : SP6"
BASE COUNT 268 a 126 c 100 g 195 t 412 others
ORIGIN
Query Match 4.2%; Score 39.4; DB 123; Length 1101;
Best Local Similarity 6.8%; Pred.No.0.3;
Matches 15; Conservative 123; Mismatches 81; Indels 0; Gaps 0;
QY 681 ccaccucgcucagcagggagcgcccgacucgagcucuguguaucgaaugugug 740
Db 856 YCYCCYBBSYBHYCYCYCCYTYTKDRRAWTTSTKBKKTTKKTKTKKSGK 915
QY 741 accgucuccaagcagcgugaaauaagucguguuuguguuuuuuuuuuuuuuuu 800
Db 916 KBKKKKSTSBKYBVGYBBYCGCTSBSSCTKTCYBCCYTCSTKTTTKYKTBTT 975
QY 801 uuugucgucacuuugucagcagcccccgaaugugugugccacaccgcgcgucg 860
Db 976 TKTTKKKTGTTGKKYBTTTYYKSBKKTGKGKGYGBYKSSGYSCSKCBSBCKBY 1035
QY 861 gaauaauuuuuugagagucuuuuuucccagugucuc 899
Db 1036 GCBCICYBKTTTTTKTKCTKTBTTTKYKSBBIYKC 1074
RESULT 6
CNS015XR
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC BACN15O17 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL106041
VERSION AL106041.1 GI:5619746
KEYWORDS GSS.

SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 1159)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES	location/Qualifiers
source	1..1159
	/organism="Drosophila melanogaster"
	/plasmid="pBelobAC11"
	/db_xref="taxon:7227"
	/clone_lib="DrosBAC"
	/clone="BACN15017"
	/note="end : T7"
BASE COUNT	448 a 36 c 7 g 178 t 490 others
ORIGIN	
Query Match	4.2%; Score 39.2; DB 123; Length 1159;
Best Local Similarity	3.7%; Pred. No.0.35;
Matches	15; Conservative 230; Mismatches 161; Indels 1; Gaps 1;
OY	530 auacacaauuggggucgucgcgggaauccucccaucgagcgaagugccuacuguuuc
Db	676 ATDCWMCYTTYSAAVAKTWTSTKAAKMBTBCKCAAVDTCDYGTTTBTSTKYTSTBS
OY	590 uucgaacuccgcgcgcgguagauguacuuuguauuuuguaaccucgaggguuuggag
Db	736 NAYRSASWBTKGMTVADKSTBKNAATAAWTSTBTBTSSTBSYSTBTBYSTDYKST
OY	650 gaucggagugggcgagcgcgcgcggaagcuccaaccucgcucacgagggagcgcucca
Db	796 BBTBWSWTTTATATATATSCSSISBSSSSTBSSTGCKSTSTSSBBABTWSTVTKT
OY	710 ucugagcucuguguaucugauuuuguaaccgucuccaagacguguaauaauaagu
Db	856 TGTSTSBCTTBTBYKSBTSTSSBBSBBSBBBBDDBKSTSTKSBSSTSTSTBST
OY	770 cguguuuuguguuuuuuuuuuuacuuuguuuuuguc-acuugucgacagcgcccu
Db	916 STTSTSTSTSTSTSTSTSTSBBSSTCTSKSKSTSKCGBSTTSTSTGSKYGSTSS
OY	829 gcgaauugugugccacacacgcgcgcgcgcuugcgaaauaauacuugagagucuuu
Db	976 SSSBSSTSSBSBBSBBSKBTSTSTGSTSTSBSTSTKSTKSTBSBBSSTSSBS
OY	889 ccagugucuccgguuuuacucgucucucuccucuccucgagcgcggg 935
Db	1036 TKTBTBTBTBTBTBTBTSTTTTSSSTBTBTBSBSBSTBBS 1082
RESULT	7
CNS016TE	
LOCUS	CNS016TE 1010 bp DNA GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN16C05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL107180
VERSION	AL107180.1 GI:5626016
KEYWORDS	GSS.

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 11:56:02 ; Search time 3871.63 Seconds
(without alignments)
658.348 Million cell updates/sec

Title: US-09-214-124-2
Perfect score: 578
Sequence: 1 ggggucgcgcgucacacau.....ccucucgcgcgggauggg 578

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues
Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: em_est35: *
83: em_est36: *
84: em_est37: *
85: gb_est48: *
86: gb_est49: *
87: gb_est50: *
88: gb_est51: *
89: gb_est52: *
90: gb_est53: *
91: gb_est54: *
92: gb_est55: *
93: gb_gss1: *
94: gb_gss2: *
95: gb_gss3: *
96: gb_gss4: *
97: em_gss1: *
98: em_gss2: *
99: em_gss3: *
100: em_gss4: *
101: gb_gss5: *
102: gb_gss6: *
103: gb_gss7: *
104: gb_gss8: *
105: gb_gss9: *
106: em_gss5: *
107: em_gss6: *
108: em_gss7: *
109: em_gss8: *
110: em_gss9: *
111: em_gss10: *
112: em_gss11: *
113: gb_gss10: *
114: gb_gss11: *
115: em_gss12: *
116: gb_gss12: *

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN			
CNS0119V/c	Drosophila melanogaster genome survey sequence T7 end of BAC BACN05C05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL099997	AL099997.1	GI:5611608	GSS.	Drosophila melanogaster	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	fruit fly.					157 a	38 c	17 g	287 t	308 others
RESULT 2	Query Match	Best Local Similarity	7.4%;	Score 42.6;	DB 122;	Length 807;											
	Matches 14;	Conservative 129;	Mismatches 97;	Indels 0;	Gaps 0;												
OY	228 uucgaacuccgcgcgcgguagauguacuuguuuuugucacgcgaggguuugag	287															
Db	763 WTTNATSTVSTSSWSTSBTSKSTAKVTDDBABRSTSTTNSSTKSSCGKGTAKKGGK	704															
OY	288 gaucggaguggcgggacgcugccgcgggaagcuccacucgcgcagcaggggagcccu	347															
Db	703 TGTGCTGTGTGCTGTSTGWSKSBGCTGKTBTBTBTDTSTSTSTNTBTBTSTSTS	644															
OY	348 ucugagcucugugguacuuguuguugagcgcugucuccaagacgguauauaagu	407															
Db	643 TBKSSSTTTSTKTSTSTSTSTBTBTKTKTKSTSTSTBTDTTTSBTRGKWGCTKG	584															
OY	408 cgugguuguguguuuguuacucuguguuugucacucacugucacugcagcagccug	467															
Db	583 KGTAKBRTGKTWGTCTTTTSTTTTSTTKTTTBTBTSTTTTSTBTSTTSBSSBS	524															
RESULT 3	Query Match	Best Local Similarity	7.4%;	Score 42.6;	DB 122;	Length 807;											
	Matches 14;	Conservative 129;	Mismatches 97;	Indels 0;	Gaps 0;												
CNS006U0	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR4N21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL065923	AL065923.1	GI:4944891	GSS.	Drosophila melanogaster	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	fruit fly.									
LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN			

[illegible]

[illegible]

JOURNAL	COMMENT	FEATURES
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.		Location/Qualifiers 1. 1159 /organism="Drosophila melanogaster" /plasmid="pBelobAC11" /db_xref="taxon:7227" /clone_lib="DrosBAC" /clone="BACN15O17" /note="end : T7"
BASE COUNT	448 a 36 c 7 g 178 t 490 others	
ORIGIN		
Query Match	6.8%; Score 39.2; DB 123; Length 1159;	
Best Local Similarity	3.7%; Pred. No. 0.38;	
Matches	15; Conservative 230; Mismatches 161; Indels 1; Gaps 1;	
QY 168	aaacacaauuggggagcucgcgggaauucccccacagcgagaagucguuuuc 227	
Db 676	ATDCWCCTTYTSAVAKTWSTKAAKBMTBCKCAAVPDCYGTTTBTSTKYTSBSsv 735	
QY 228	uucgaacuccgcgcgcgcuuaaguacuuguuuuugcuaccucgcgaggguuuggag 287	
Db 736	NAYRSASMBTKGMTVADKSGBKMATAAAMTSTBTBSSSTBBStSTBTBYSTDITYKST 795	
QY 288	gaucggaguggcgggacgcgcgcgggaagcuccaccucgcgcucagcaggagcccuca 347	
Db 796	BByTTBMSWTTTATTATSTAATSCSSSTSBSSSSSTBSITGCKSTSTSTSSBABTBWTYTKT 855	
QY 348	ucugagcucuguguaucugaauuguugagaccgucuccaagaagcgugauanaaagu 407	
Db 856	TSTSTBCCTBTBYKSBSTSSBSBSTSBBSBSTDBBKSTSTKSBSSTSTBTBSTST 915	
QY 408	cguuguuuuguguuuuuuguaucuuguguguuuguc-acuugucagacgcgccu 466	
Db 916	STTSTSTSTSTSTSTSTSTSBBSSTCTSKSKSTSKCGBSTTSTTSTGSKYGSTSSSBS 975	
QY 467	gcgaauugugugccccacacgcgcgcgcgcgcgaauaauacuugagagagucuuugcu 526	
Db 976	SSSBSTSSSBSBBSKTBTSTGSTSTSBSTSTKSTBSBBTSSBSBTSTSTSTIKBT 1035	
QY 527	ccagugucuccguuuuguaucugucuccucuccucuccgcgcggg 573	
Db 1036	TKBTBTBTBTBTBTBTBTSTTTTSTSSSTBTBTBSBSBTSTBTBBS 1082	
RESULT 6		
CNS0166K		
LOCUS	CNS0166K 1225 bp DNA GSS 26-JUL-1999	
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL106358	
VERSION	ALI06358.1 GI:5621512	
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1225)	
REFERENCE	Genoscope. Direct Submission	

JOURNAL
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebl.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAc11.
COMMENT
FEATURES
source Location/Qualifiers
CNS016TE 1..1225 /organism="Drosophila melanogaster"
LOCUS DEFINITION /plasmid="pbelOBac11"
BACN16C05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence. /db_xref="taxon:7227"
AL107180 /clone_lib="DrosBAC"
GI:5626016 /note="end : SP6"
BASE COUNT 84 a 138 c 318 g 290 t 395 others
ORIGIN
Query Match 6.7%; Score 38.6; DB 123; Length 1225; Best Local Similarity 15.4% Pred. No. 0.57;
Matches 74; Conservative 165; Mismatches 235; Indels 5; Gaps 1;
OY 1 ggaggcccgccgucuaacacaugaugagcgcggcccagaauccgaucguaaanaag 60
Db 667 GKkKtTGGCCCTTTTKTkTktTBbbBSBBSSSBcssSKBYTYCGGGRWtt 726
OY 61 uuuuuuccuaaunaucucagaauggcagugagagagaaaauuguugugugaucu 120
Db 727 TWTMTTGTTKTkrMTTTTTTtkTKKGggGGGDgdGGGGGTWtgTtGKgGgtTTTTtt 786
OY 121 gccucaucggugggnaaggguaccgagcaucguaucguuuucgauaacaaauug 180
Db 787 TGtTTTKTTTKGgTTgGgkgKGggGGGGGGGGGGGGGGGGGGGtKGT----KTKKGT 841
OY 181 ggguccgucggggaauccccccaucggcagaaugucuaucguuuucuugaacuccgc 240
Db 842 GTGTGCgTTtGTTGWtTTGKKkkGTkKGGKKGKKGKKGKKGKkKtTGKKGKKtGGGK 901
OY 241 gcgguaagaaaguacuugauuuuguaaccucgcagcgagggagcuuuuggaagcaucggagcg 300
Db 902 KKGCGGKkkKKKKKKGtGGGGGtKKKKGGGGTtGGKKKKGTGKrGTGKGGGGGGGK 961
OY 301 ggaacgucgccgggaagcacuccaccucgcucagcagggagcgcccgauucugagcucug 360
Db 962 GTGctTGKGKGtGKGTtTGcGgTGGGtTtTGtGGKKGggGtGTtTKGWKGTGtGKt 1021
OY 361 guaacugaauugaugagcagcucaccaagacguguaaauaagaucguuguugug 420
Db 1022 TTATATTGKtKtTGTGKGKKGKGTGKGKtGKDgKkKtTKKGGGtTKGtGtGKtG 1081
OY 421 uuuguuuuuuacuuuguuuuugucacuuugucagacagcccucgcaaauugugug 479
Db 1082 KGtGTWtGTGtGtKKKKGGTgTGGGtGtGTkTKTGKGGGGtGTGtGtGtGK 1140
RESULT 7
CNS016TE 1010 bp DNA GSS 26-JUL-1999
LOCUS DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN16C05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION ALI07180
VERSION ALI07180.1 GI:5626016
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
1	(bases 1 to 1010)	Genoscope.	Direct Submission	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	Location/Qualifiers
				Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.	1. .1010
					/organism="Drosophila melanogaster"
					/plasmid="pBelobAC11"
					/db_xref="taxon:7227"
					/clone_lib="DrosBAC"
					/clone="BACN16C05"
					/note="end : SP6"
BASE COUNT	75 a	191 c	154 g	134 t	456 others
ORIGIN					

[illegible]

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .910
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1lb="RPCI-98"
/clone="BACR14J21"
/note="end : T7"

BASE COUNT 202 a 63 c 112 g 198 t 335 others
ORIGIN

Query Match 6.4%; Score 36.8; DB 122; Length 910;
Best Local Similarity 3.4%; Pred. No. 1.9;
Matches 10; Conservative 178; Mismatches 102; Indels 2; Gaps 2;

QY 269 ccucgcgagguuugggaggaucgagugcgagcgagccgcgggaagccaccuccgc 328
Db 620 SSTSGCGYSSSGSBSSCSSTBTSSSTBTBTKBTSTYSTSTTTSTTBSBTSYGS 679
QY 329 ucagcagggagcgccgcaucgagcucguguaucgaaucgugugagccgucucca 388
Db 680 BB-KSTSSSTSSSTSSSSSSSTSSCTSTKBYSTSTBSYBTTBTBTSTBTBTST 738
QY 389 agacggugauaauaagucguguuuguguuuguuuuaaccuuguuugucuc 448
Db 739 BTBTBTBTBTSTSTMTSTYTBTBTBTSTBTSTBTSTBTSTBTSTBTSTBTST 798
QY 449 acuuugcgcagc 508
Db 799 SYSTSTSCBS-SBSBSSTSYSBCTSTSTSTSSBSBSBSBSBSBSBSBSBSBS 857
QY 509 uuggagagucuuuugcuccagucucucgucuuugucucucucucucucc 560
Db 858 TTCTKTGTGBSSYGTGTSTYTTBTATTSTSTSTSTSSBSBTSTBSBTSTBS 909

RESULT 13

CNS00JX7 977 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR39C01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL076850
VERSION AL076850.1 GI:4956428
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 977)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
ECORI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's

p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .977
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1lb="RPCI-98"
/clone="BACR39C01"
/note="end : TET3"

BASE COUNT 132 a 80 c 110 g 263 t 392 others
ORIGIN

Query Match 6.4%; Score 36.8; DB 122; Length 977;
Best Local Similarity 10.3%; Pred. No. 1.9;
Matches 36; Conservative 134; Mismatches 177; Indels 1; Gaps 1;

QY 79 ucagauggcagugagaggaauuuugucgugugagcgcgcgcgcgcgcgcgcgcgcgc 138
Db 4 TCCTATGGCGCGGGGGGCTCAGTTTGKRGGGGGCGGGCGGAGTGGGKGKG 63
QY 139 ggggucgc 198
Db 64 GGGGGCTTTGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGG 123
QY 199 ucccaucgcagagugcucucucucucucucgcgcgcgcgcgcgcgcgcgcgcgc 258
Db 124 NNN 183
QY 259 gaauugucucgc 318
Db 184 SSSSTTSSATSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 242
QY 319 ccaccucgc 378
Db 243 SSSBSBSASBS 302
QY 379 accgucuccaagc 426
Db 303 STTSSBSBTNTVSTSTTTATTASGMTSATSTTWTSTGSAVATWTTT 350

RESULT 14

CNS016LW 1101 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACN16J16 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL106910
VERSION AL106910.1 GI:5624430
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES

Location/Qualifiers

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:28:16 ; Search time 172.61 Seconds
(without alignments)
460.618 Million cell updates/sec

Title: US-09-214-124-2
Perfect score: 578
Sequence: 1 ggggucgcgcgcucacacau.....ccucucgcgcgcgcgaugg 578

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5C_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/5D_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/6_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
7: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.6	6.5	7218	1	US-08-232-463-14 Sequence 14, Appl
2	30.6	5.3	1276	4	US-08-793-599-1 Sequence 1, Appl1
3	30.6	5.3	2142	4	US-08-793-599-3 Sequence 3, Appl1
4	29	5.0	13158	4	US-08-687-080-105 Sequence 105, App
5	28.8	5.0	3934	5	US-09-226-568-18 Sequence 18, Appl
6	28.8	5.0	3946	1	US-08-077-848A-1 Sequence 1, Appl1
7	28.8	5.0	3946	5	US-09-211-640-1 Sequence 1, Appl1
8	28.8	5.0	3946	6	PCT-US94-03547-1 Sequence 1, Appl1
9	28.8	5.0	11531	1	US-08-068-945A-1 Sequence 1, Appl1
10	28.8	5.0	11531	2	US-08-442-806-1 Sequence 1, Appl1
11	28.2	4.9	221	1	US-08-222-177A-17 Sequence 17, Appl
12	28.2	4.9	278	1	US-08-222-177A-52 Sequence 52, Appl
13	28.2	4.9	9636	2	US-08-323-170B-1 Sequence 1, Appl1
14	28	4.8	3240	2	US-08-162-081B-34 Sequence 34, Appl
15	28	4.8	3240	3	US-08-780-872-34 Sequence 34, Appl
16	28	4.8	3412	2	US-08-162-081B-32 Sequence 32, Appl
17	28	4.8	3412	3	US-08-780-872-32 Sequence 32, Appl
18	27.8	4.8	3566	5	US-08-689-421-32 Sequence 32, Appl
19	27.6	4.8	6545	4	US-08-843-530B-1 Sequence 32, Appl
20	27.6	4.8	8224	4	US-09-010-398-14 Sequence 1, Appl1
21	27.4	4.7	943	5	US-08-705-875A-1 Sequence 14, Appl
22	27.2	4.7	2055	1	US-07-842-349-1 Sequence 1, Appl1
23	27.2	4.7	2277	2	US-08-676-967-2 Sequence 2, Appl1
24	27.2	4.7	2277	2	US-08-676-967-2 Sequence 2, Appl1
25	27.2	4.7	2277	4	US-09-098-487-2 Sequence 2, Appl1
26	27.2	4.7	2643	4	US-08-781-802-11 Sequence 11, Appl

C 27	27.2	4.7	3255	3	US-08-916-917-11	Sequence 11, Appl
C 28	27.2	4.7	3255	5	US-09-225-170-11	Sequence 11, Appl
C 29	27.2	4.7	3545	4	US-08-781-802-9	Sequence 9, Appl1
C 30	27.2	4.7	4315	4	US-08-781-802-1	Sequence 1, Appl1
C 31	27.2	4.7	7208	5	US-09-166-186-107	Sequence 107, App
C 32	27	4.7	2281	2	US-08-164-614A-1	Sequence 1, Appl1
C 33	27	4.7	2281	4	US-08-456-489B-1	Sequence 1, Appl1
C 34	27	4.7	2281	6	PCT-US93-01720-1	Sequence 1, Appl1
C 35	27	4.7	2488	2	US-08-279-270A-2	Sequence 2, Appl1
C 36	26.8	4.6	20235	2	US-07-642-734C-3	Sequence 3, Appl1
C 37	26.4	4.6	358	2	US-08-254-404-19	Sequence 19, Appl
C 38	26.4	4.6	358	4	US-08-327-451E-19	Sequence 19, Appl
C 39	26.4	4.6	358	4	US-08-458-109-19	Sequence 19, Appl
C 40	26.4	4.6	741	5	US-08-937-271-12	Sequence 12, Appl
C 41	26.4	4.6	1029	5	US-08-937-271-16	Sequence 16, Appl
C 42	26.4	4.6	1729	3	US-08-844-024-1	Sequence 1, Appl1
C 43	26.4	4.6	1729	4	US-08-718-547-1	Sequence 1, Appl1
C 44	26.2	4.5	500	5	US-09-141-000-2	Sequence 2, Appl1
C 45	26.2	4.5	1155	5	US-08-860-368B-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-F1s
US-08-232-463-14


```

;      ANTI-SENSE: NO
;      ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE: 5' END OF INTRON 21 OF RAD50 GENOMIC
;      INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-105

```

Query Match	5.0%;	Score 29;	DB 4;	Length 13158;
Best Local Similarity	34.0%;	Pred. No. 13;		
Matches	18;	Conservative	20;	Mismatches 15;
				Indels 0;
				Gaps 0;

[illegible]

```

RESULT      5
US-09-226-568-18/c
; Sequence 18, Application US/09226568
; Patent No. 6001992
;
; GENERAL INFORMATION:
;
; APPLICANT: Ackermann, Elizabeth J.
;
; APPLICANT: Bennett, C. Frank
;
; APPLICANT: Dean, Nicholas M.
;
; APPLICANT: Marcusson, Eric G.
;
; TITLE OF INVENTION: Antisense Modulation of No. 6001992el Anti-apoptotic
; FILE REFERENCE: ISPH-0337
;
; CURRENT APPLICATION NUMBER: US/09/226,568
;
; CURRENT FILING DATE: 1999-01-07
;
; NUMBER OF SEQ ID NOS: 39
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 18
;
; LENGTH: 3934
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence: antisense
;
US-09-226-568-18

```

Query Match	5.08;	Score 28.8;	DB 5;	Length 3934;
Best Local Similarity	56.28;	Pred. No. 8.1;		
Matches 27; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0

Qy	525	cuccagugucuuccgguuuagucgucucucuccucucgcgcgcg	572
Db	600	1: : : : : : : : ccccccgactgccgggtacaaactctgctctctctctctgctggcgg	553

```

RESULT 6
US-08-077-848A-1/c
: Sequence 1, Application US/08077848A
: Patent No. 5470955
: GENERAL INFORMATION:
: APPLICANT: Craig, Ruth W.
: TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
: TITLE OF INVENTION: POLYPEPTIDE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 1880 Century Park East, Suite 500
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90067
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:

```

```

: APPLICATION NUMBER: US/08/077,848A
: FILING DATE: 16-JUN-1993
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Hallé, Ph.D., Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: PD-2845
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 455-5100
: TELEFAX: (619) 455-5110
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3946 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: mcl-1
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 61..1110
: OTHER INFORMATION: /note="When nucleotide 740 = C,
: OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
: OTHER INFORMATION: acid 227 = V."
US-08-077-848A-1

```

Query Match	5.08;	Score 28.8;	DB 1;	Length 3946;
Best Local Similarity	56.28;	Pred. No. 8.1;		
Matches 27; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0;

Qy 525 cuccagugucuccguuuuagacucgucucucucuccucucgcgg 572
|:|:| | : |:| : |:|:|:|:| |:| : | |
Db 600 CTCACGCACTGCCGGTACAATCGTCCTCCTCCCTGTGTCGGCG 553

RESULT 7
US-09-211-640-1/c
; Sequence 1, Application US/09211640
; Patent No. 6020466
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,640
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/441,375
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note="When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
US-09-211-640-1

Query Match          5.0%; Score 28.8; DB 5; Length 3946;
Best Local Similarity 56.2%; Pred. No. 8.1;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 525 cuccagugucuuuccguuuuacucgucuccucuccucgcccgcg 572
      ||||| | | | | | | | | | | | | | | | | | | | |
Db 600 CTCACGCGACTGCCGGTACACTCGTCTCTCTCTCTGTCGCGG 553

RESULT 8
PCT-US94-03547-1/c
; Sequence 1, Application PC/TUS9403547
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: MYELOID CELL LEUKEMIA ASSOCIATED GENE
; TITLE OF INVENTION: MCL-1
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03547
; FILING DATE: 31-MAR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mcl-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1110
; OTHER INFORMATION: /note="When nucleotide 740 = C,
; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
; OTHER INFORMATION: acid 227 = V."
PCT-US94-03547-1
```

```

Query Match          5.0%; Score 28.8; DB 6; Length 3946;
Best Local Similarity 56.2%; Pred. No. 8.1;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 525 cuccagugucuuuccguuuuacucgucuccucuccucgcccgcg 572
      ||||| | | | | | | | | | | | | | | | | | | | |
Db 600 CTCACGCGACTGCCGGTACACTCGTCTCTCTCTCTGTCGCGG 553

RESULT 9
US-08-068-945A-1
; Sequence 1, Application US/08068945A
; Patent No. 5616483
; GENERAL INFORMATION:
; APPLICANT: Bjursell, Gunnar
; APPLICANT: Carlsson, Peter
; APPLICANT: Enerback, Sven
; APPLICANT: Hansson, Lennart
; APPLICANT: Lidberg, Ulf
; APPLICANT: Nilsson, Jeanette
; APPLICANT: Tornell, Jan
; TITLE OF INVENTION: New DNA Sequences
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,945A
; FILING DATE: 27-MAY-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9201809-2
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9201826-6
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9202088-2
; FILING DATE: 03-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9300902-5
; FILING DATE: 19-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)819-8783
; TELEFAX: (212)354-8113
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Mammary gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
```



```

? LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
? LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
? LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
? LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
? OTHER INFORMATION: /EC_number-3.1.1.1
? OTHER INFORMATION: /product- "Bile Salt-stimulated lipase"
? FEATURE:
? NAME/KEY: 5'UTR
? LOCATION: 1..1640
? FEATURE:
? NAME/KEY: TATA_signal
? LOCATION: 1611..1617
? FEATURE:
? NAME/KEY: exon
? LOCATION: 1641..1727
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4071..4221
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4307..4429
? FEATURE:
? NAME/KEY: exon
? LOCATION: 4707..4904
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6193..6323
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6501..6608
? FEATURE:
? NAME/KEY: exon
? LOCATION: 6751..6868
? FEATURE:
? NAME/KEY: exon
? LOCATION: 8335..8521
? FEATURE:
? NAME/KEY: exon
? LOCATION: 8719..8922
? FEATURE:
? NAME/KEY: exon
? LOCATION: 10124..10321
? FEATURE:
? NAME/KEY: exon
? LOCATION: 10650..11490
? FEATURE:
? NAME/KEY: 3'UTR
? LOCATION: 11491..11531
?
US-08-945A-1

```

[illegible]

```

: Sequence 1, Application US/08442806
: Patent No. 5716817
:
: GENERAL INFORMATION:
: APPLICANT: Bjursell, Gunnar
: APPLICANT: Carlsson, Peter
: APPLICANT: Enerback, Sven
: APPLICANT: Hansson, Lennart
: APPLICANT: Lidberg, Ulf
: APPLICANT: Nilsson, Jeanette
: APPLICANT: Tornell, Jan
: TITLE OF INVENTION: Genomic DNA Sequences
: TITLE OF INVENTION: Encoding Human BSSL/CEL
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: White & Case
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2787
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/442,806
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/068,945
: FILING DATE: 27-MAY-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9201809-2
: FILING DATE: 11-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9201826-6
: FILING DATE: 12-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9202088-2
: FILING DATE: 03-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9300902-5
: FILING DATE: 19-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Sterner, Richard J.
: REGISTRATION NUMBER: 35,372
: REFERENCE/DOCKET NUMBER: 1103326-052
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)819-8783
: TELEFAX: (212)354-8113
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11531 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: TISSUE TYPE: Mammary gland
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
: LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
: LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
: LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
: LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
: OTHER INFORMATION: /EC_number= 3.1.1.1

```


Sequence 1, Application US/08323170B
Patent No. 5733772
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaslow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf5230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-323-170B-1

```

Query Match          4.9%; Score 28.2; DB 2; Length 9636;
Best Local Similarity 31.5%; Pred. No. 21;
Matches 28; Conservative 23; Mismatches 38; Indels 0; Gaps 0;

Oy 34 ggcgcagaucgaaucuguaaaguuuuuuucucuaaucucagauggcaguga 93
      |||||:|::|||:::|:::|:::|:::|:::|:::|
Db 4797 GGTCTAATGGTGAATTGGACATTATATTTTTT TTTT TTTTAAATCAACACATATACATTTA 4738

Oy 94 gagagauuuugucugugugagcgug 122
      ||||:::|:::|
Db 4737 AATGATATATCATTTATTCATTAAAGCTGG 4709

RESULT 14
US-08-162-081B-34
; Sequence 34, Application US/08162081B
; Patent No. 5824492
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
;

```

```

1 ADDRESSSEE: Felfe & Lynch
2 STREET: 805 Third Avenue
3 CITY: New York
4 STATE: New York
5 COUNTRY: USA
6 ZIP: 10022
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
10
11 COMPUTER: IBM PS/2
12 OPERATING SYSTEM: PC-DOS
13 SOFTWARE: Wordperfect
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/162,081B
17 FILING DATE: February 7, 1994
18
19 CLASSIFICATION: 435
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: PCT/GB93/00761
23 FILING DATE: 13 April 1993
24
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Pasqualini, Patricia A.
27 REGISTRATION NUMBER: 34,894
28 REFERENCE/DOCKET NUMBER: LUD 5256
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (212) 688-9200
31 TELEFAX: (212) 838-3884
32
33 INFORMATION FOR SEQ ID NO: 34:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 3240 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39
40 US-08-162-081B-34

```

	Query Match	4.88;	Score 28;	DB 2;	Length 3240;	
	Best Local Similarity	34.2%;	Pred. No. 13;			
	Matches	26;	Conservative	20;	Mismatches	30; Indels 0; Gaps 0;
QY	404	aagucgugguuuuguguuuacccuuguguuugucgcacucgacagcgc	463			
		: : : : : : : : :				
Dd	2475	AAATCAAGGCTTGATCCTTCGAATGTACCTTATGGTTGTCTGTCCAATCGGTGACTGTGT	2534			
QY	464	ccugcgaaauugugug	479			
		- : : :				
Dd	2535	GGGACTTATTGAGGTG	2550			

RESULT 15
US-08-780-872-34
; Sequence 34, Application US/08780872
; Patent No. 5846824
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780, 872

Tue Aug 8 08:33:19 2000

us-09-214-124-2.rn1

Page 9

```

1      FILING DATE: 09-JAN-1997
2      CLASSIFICATION: 435
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: 08/162,081
5      FILING DATE: February 7, 1994
6      APPLICATION NUMBER: PCT/GB93/00761L
7      FILING DATE: 13 April 1993
8      ATTORNEY/AGENT INFORMATION:
9      NAME: Pasqualini, Patricia A.
10     REGISTRATION NUMBER: 34,894
11     REFERENCE/DOCKET NUMBER: LUD 5256
12     TELECOMMUNICATION INFORMATION:
13     TELEPHONE: (212) 688-9200
14     TELEFAX: (212) 838-3884
15     INFORMATION FOR SEQ ID NO: 34:
16     SEQUENCE CHARACTERISTICS:
17     LENGTH: 3240 base pairs
18     TYPE: nucleic acid
19     STRANDEDNESS: single
20     TOPOLOGY: linear
21
22 US-08-780-872-34

```

```

Query Match          4.8%; Score 28; DB 3; Length 3240;
Best Local Similarity 34.2%; Pred. No. 13;
Matches 26; Conservative 20; Mismatches 30; Indels 0; Gaps 0;

QY 404 aaugcuguguuuguguuuuuuuacuuuguguuugucgacuuugcagcgc 463
    ||::||::||::||::||::||::||::||::||::||::||::||::||
Db 2475 AAATCAAGGCTCTTGATCTTCGAATGTTACCTTATGTTGTCTGTCAATCGGTGACTGTGT 2534
    ||::||::||::||::||::||::||::||::||::||::||::||

QY 464 ccugcgaaugugugug 479
    ||::||::||::||::||::||::||::||::||::||::||::||
Db 2535 GGGACTTATTTGAGGTG 2550
    ||::||::||::||::||::||::||::||::||::||::||::||

```

Search completed: August 7, 2000, 09:28:43
Job time: 897 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 10:53:33 ; Search time 4994.97 Seconds
(without alignments)
1200.515 Million cell updates/sec

Title: US-09-214-124-2
Perfect score: 578
Sequence: 1 ggggucgcccgcucacacau.....ccucucgcccgggauggg 578

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 5187315402 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl: *
1: gb_bal: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_pat: *
6: gb-ph: *
7: gb-pl1: *
8: gb-pl2: *
9: gb-pl1: *
10: gb-pr2: *
11: gb-pr3: *
12: gb-ro: *
13: gb-sts: *
14: gb-sy: *
15: gb-un: *
16: em_fun: *
17: em_hum1: *
18: em_hum2: *
19: em_in: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em-ph: *
25: em-pl: *
26: em-ro: *
27: em-sts: *
28: em-sy: *
29: em-un: *
30: em-v1: *
31: gb_htg1: *
32: gb_htg2: *
33: gb_in1: *
34: gb_in2: *
35: em_bal: *
36: em_ba2: *
37: em_hum3: *
38: em_hum4: *
39: gb-pr4: *
40: gb_htg3: *
41: gb_htg4: *
42: gb_htg5: *
43: gb_htg6: *

44: gb_htg7: *
45: em_htg1: *
46: em_htg2: *
47: em_htg3: *
48: em_hum5: *
49: gb-pl3: *
50: gb-pr5: *
51: gb_htg8: *
52: gb_htg9: *
53: gb_htg10: *
54: gb_htg11: *
55: gb_htg12: *
56: gb_htg13: *
57: gb_htg14: *
58: gb_in3: *
59: gb_htg15: *
60: gb_htg16: *
61: gb_htg17: *
62: em_htg4: *
63: em_htg5: *
64: em_htg6: *
65: em_htg7: *
66: em_hum6: *
67: gb_htg18: *
68: gb_htg19: *
69: gb_htg20: *
70: gb_htg21: *
71: gb_htg22: *
72: gb_htg23: *
73: gb_htg24: *
74: gb_htg25: *
75: gb_htg26: *
76: gb_htg27: *
77: gb_htg28: *
78: gb_htg29: *
79: gb_htg30: *
80: gb_htg31: *
81: gb_v11: *
82: gb_v12: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	578	100.0	578	5 A83549	A83549 Sequence 2
2	578	100.0	940	5 A83548	A83548 Sequence 1
3	546.8	94.6	4292	82 AF006065	AF006065 Fowlpox v
4	423.6	73.3	1530	81 RESNVX	V01200 Spleen necr
5	310	53.6	891	81 REXXX1	V01204 Spleen necr
6	173.4	30.0	1005	81 ACRLTR1	M22223 Reticuloend
7	171	29.6	545	81 S70398	S70398 {LTR, U3, R
8	156.4	27.1	585	81 S82226	S82226 {RS region,
9	153	26.5	887	81 S79845	S79845 (REV LTR) (
10	150	26.0	887	81 ACRLTR2	M22224 Reticuloend
11	121	20.9	859	81 REXXX2	M12205 Spleen necr
12	113.6	19.7	160	4 CHKSNTIE2	M12248 Spleen necr
13	113	19.6	180	81 SNVTLR	X59450 Spleen necr
14	111.6	19.3	160	4 CHKSNTIB2	M12242 Spleen necr
15	111	19.2	160	4 CHKSNTID2	M12246 Spleen necr
16	111	19.2	160	4 CHKSNTIF2	M12250 Spleen necr
17	110.6	19.1	160	4 CHKSNTIA2	M12206 Spleen necr
18	110.6	19.1	160	4 CHKSNTIC2	M12244 Spleen necr
19	44.8	7.8	290	13 AV025522	AV025522 Rattus no
20	38.2	6.6	1143	4 BB074351	U74351 Bulweria bu
21	37.6	6.5	378	13 AV026236	AV026236 Rattus no
22	37.6	6.5	7218	5 I66494	I66494 Sequence 14
23	37.2	6.4	113440	55 AC020935	AC020935 Homo sapi
24	37	6.4	70480	39 AC006973	AC006973 Homo sapi

25	37	6.4	173618	57	AC019045
26	36.4	6.3	1143	4	AF076053
27	36.2	6.3	2041	12	MUSENZE1A
28	36	6.2	16832	42	AC014255
29	36	6.2	88866	41	AC010696
30	36	6.2	114958	51	AC008356
31	36	6.2	119182	51	AC008234
32	36	6.2	224230	34	AE003682
33	35.8	6.2	151820	40	AL157931
34	35.8	6.2	186518	60	AC016813
35	35.6	6.2	1026	4	MTPACYT21
36	35.4	6.1	72394	75	AC040913
37	35.2	6.1	176043	77	AC027072
38	35.2	6.1	213721	10	HS172B20
39	35.2	6.1	346558	77	AC021061
40	35	6.1	155278	10	HSDJ655C4
41	34.8	6.0	168613	40	AL139318
42	34.6	6.0	1084	31	FH059841
43	34.6	6.0	167656	4	AP001645
44	34.4	6.0	68734	51	AC023218
45	34.4	6.0	92958	31	DMBR11J17

Db	241	GCCGCTAAGTAAGTACTTGAITTTGGTACCTCGCGAGGGTTTGGAGATCGGAGTGGCG	300
QY	301	ggacgcugccgggaaagcuccacccgcgcucacgacgagggagcccgauaucgagcucugug	360
Db	301	GGACGCTGCGCGGAAAGCTCCACCTCCGCTCAGCAGGGGACGCCCTGATCTGAGCTCTGTG	360
QY	361	guaucugauugugugugagaccgucuccaagaacgugugauaanaaagucugugugugugug	420
Db	361	GTATCTGATGTGTGTGGACCGCTCTCCAGACGGGTGATATATAAGTCGTGTTGTGTG	420
QY	421	uuuguuuuguaaccuuguguuugucgucacuuugcagacgcgcccgcgaaauugugugc	480
Db	421	TTTGTTGTGTACCTGTGTGTTGTTGCTGCTACTTGTGCACAGCGCCCTGCGAATTGTTGTGC	480
QY	481	ccacacccgcgcgcgucugcgaaauaacuuugagagucuuugccuccagugucuccgu	540
Db	481	CCACACCGCGCGGCTTCGGAATAATACTTTGGAGAGTCTTTTGGCTTCAGTGTCTCCGT	540
QY	541	uuguaucugucuccucuccuccucgcgcgcggauggg	578
Db	541	TTGTACTCGTCTCCTCCTCCTCCTCCTCGGCGCGGATGGG	578

		ALIGNMENTS					
RESULT	1						
A83549		578 bp	DNA	PAT		21-JAN-2000	
LOCUS							
DEFINITION		Sequence 2 from Patent WO9849334.					
ACCESSION		A83549					
VERSION		A83549.1	GI:6732808				
KEYWORDS							
SOURCE							
ORGANISM							
		Reticuloendotheliosis virus. Reticuloendotheliosis virus Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 3-Reticuloendotheliosis virus group.					
REFERENCE							
AUTHORS		1 (bases 1 to 578) Gabus-Darlux,C. and Darlix,J. NOVEL INTERNAL RIBOSOME ENTRY SITE AND VECTOR CONTAINING SAME Patent: WO 9849334-A 05-NOV-1998; GABUS DARLIX CAROLINE (FR); INST NAT SANTE RECH MED (FR)					
TITLE							
JOURNAL							
FEATURES							
source		1..578 /organism="Reticuloendotheliosis virus" /strain="TYPE A (REV-A)" /db_xref="taxon:11636"					
BASE COUNT		94 a	139 c	172 g	173 t		
ORIGIN							
Query Match		100.0%;	Score 578;	DB 5;	Length 578;		
Best Local Similarity		70.1%;	Pred. No. 1.5e-163;				
Matches	405;	Conservative	173;	Mismatches	0;	Indels	0;
Gaps							0;
QY	1	gggugcgccgucacacaauungugagcgcgcgccagaucgaucguuaaaag	60				
Db	1	GGGGTCGCCGTCTACACATTGTTGTGAGCGCGCGCCAGATTGCAATCTGTATAAAG	60				
QY	61	uuuuuuuucuanaucucagauuggcagugagagagauuuugugugugagcu	120				
Db	61	TTTTTCTTCTATATCCTCAGATTGGCAGTGAGAGAGATTGTTGTCGTGTAGGCT	120				
QY	121	ggccuacugggugggguagggggucgcgacugaauccguaguauuucgaacaauuug	180				
Db	121	GGCCTACTGGGTGGGTAGGGGTCGGGACTGAATCCGTAGTATTTCGATACAACATTGG	180				
QY	181	gggcugcugccgggaucucccaucgcgagaagugccuacuguuucugaacucgcgc	240				
Db	181	GGGCTGCTCCGGGATTCCTCCCAATCGGCAAGAAGTGCCCTACTGTTCTTGCAGACTCCGGC	240				
QY	241	gcccguaguaagguacuuguuuuguaaccucgcgagagguuuugggaggaucggaguggcg	300				

[illegible]

[illegible][illegible]


```
REFERENCE 2 (bases 1 to 887)
AUTHORS Swift,R.A., Boerkoel,C.F., Ridgway,A., Fujita,D.J., Dodgson,J.B.
TITLE B-lymphoma induction by reticuloendotheliosis virus:
Characterization of a mutated chicken syncytial virus provirus
involved in c-myc activation
JOURNAL J. Virol. 61, 2084-2090 (1987)
MEDLINE 87226389
FEATURES
    source
        Location/Qualifiers
            1..887
                /organism="Reticuloendotheliosis virus"
                /db_xref="taxon:11636"
                246..750
                    /note="3' LTR"
            805..806
                /note="cryptic intron splice donor site"
BASE COUNT 212 a 216 c 227 g 232 t
ORIGIN About 700 bp after segment 1.

Query Match 26.0%; Score 150; DB 81; Length 887;
Best Local Similarity 64.0%; Pred. No. 1.9e-34;
Matches 114; Conservative 53; Mismatches 10; Indels 1; Gaps 1;

QY 1 ggggucgcgcguccuacacaaungugagcgcgcccgagauucgaucguuaaagaag 60
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 582 GGGGTGCGCCGTCTACACATTGTGTGACGTCGCCGCCGATTCGAATCTGTAATAAAG 641
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 uuuuuucuuuacuuuacuuuagcagauugcgagagagagauuuuugucguugaggu 120
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 642 -CTTTTCTCTATATCTCAGATTCGACGTGAGAGAGATTGTTGTCGTGCTGCTG 700
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 ggcacucuggggggggaggggucgcgacgacgaucggaucguuuuucgaucacaaauu 178
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 701 GGCCCTACTGGGTGGGGTAGGATCCGACTGCAATCCGTAGTATTTCGCTAATATATGT 758
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 11
REXXX2
LOCUS REXXX2 859 bp DNA VRL 06-APR-1993
DEFINITION Spleen necrosis virus (3' end) integrated in chicken cell.
ACCESSION V01205 J02389
VERSION V01205.1 GI:61794
KEYWORDS
SOURCE
    ORGANISM Spleen necrosis virus.
    Spleen necrosis virus
    Viruses; Retroid viruses; Retroviridae; Mammalian type C
    retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 859)
AUTHORS Shimotohno,K., Mizutani,S. and Temin,H.M.
TITLE Sequence of retrovirus provirus resembles that of bacterial
transposable elements
JOURNAL Nature 285 (5766), 550-554 (1980)
MEDLINE 80254544
FEATURES
    source
        Location/Qualifiers
            1..859
                /organism="Spleen necrosis virus"
                /db_xref="taxon:11836"
                <1..740
                    /organism="Spleen necrosis virus"
                    /proviral
                /db_xref="taxon:11836"
                741..>859
                    /note="cellular DNA"
                    /organism="Spleen necrosis virus"
                    /db_xref="taxon:11836"
BASE COUNT 245 a 199 c 214 g 201 t
ORIGIN
```

Query Match 20.9%; Score 121; DB 81; Length 859;
Best Local Similarity 64.2%; Pred. No. 1e-25;
Matches 115; Conservative 48; Mismatches 10; Indels 6; Gaps 3;

```
QY 1 ggggucgcgcguccuacaca---uuugugagcgcgcccgccagauucgaucguuaa 57
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 554 GGGGTGCGCCGTCTGACATTGTTGTTGTGACGTGCGGCCAGATTCGAATCTGTAATAA 613
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 58 aaguuuuuuucuuuacuuuacuuuagcagauugcgagagagagauuuuugucguugag 117
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 614 AAC--TTTTCTCTGAATCCTCAGATTGGCAGTGAGAGAGATTGTTGCTGCTGTG 671
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 118 gcuggcucacugggggg-guagggucgcgacugacugaaucgguuauucgaucaca 175
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 672 GCTGGCTACTGGGTGGGGCGCAGGATCCGACTGAATCCGTAGTACTTCGGTACAACA 730
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
CHKSNVIE2
LOCUS CHKSNVIE2 160 bp DNA VRT 28-APR-1993
DEFINITION Spleen necrosis virus proviral DNA/chicken DNA (clone 70), 3'
junction.
ACCESSION M12248
VERSION M12248.1 GI:212691
KEYWORDS integration site; proviral gene.
SEGMENT 2 of 2
SOURCE
    ORGANISM Chicken DNA and proviral spleen necrosis virus DNA.
    Gallus gallus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
    Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
    1 (bases 1 to 160)
    Shimotohno,K. and Temin,H.M.
    Evolution of retroviruses from cellular movable genetic elements
    Cold Spring Harb. Symp. Quant. Biol. 45, 719-730 (1981)
    81259597
FEATURES
    source
        Location/Qualifiers
            1..160
                /organism="Gallus gallus"
                /db_xref="taxon:9031"
BASE COUNT 35 a 31 c 47 g 47 t
ORIGIN About 4.5 kb downstream of segment 1.

Query Match 19.7%; Score 113.6; DB 4; Length 160;
Best Local Similarity 63.6%; Pred. No. 1.8e-23;
Matches 98; Conservative 44; Mismatches 9; Indels 3; Gaps 2;

QY 26 ugaacgcgcgcgcagaaucgaaucguuaaagaauuuuuuucuuuacuuuacuuu 85
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 1 TGACGTGCGGCCAGATTCGAATCTGTAATAAACT--TTTCTCTGAATCTCAGATT 58
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 86 ggcagugagagagauuuuugucguugagcgcgccuacugggggg-guagggguc 144
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 59 GGCAGTGAGAGAGATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 145 cggacugaaucgguuauucgaucacaaauu 178
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 119 CGGACTGAATCCGTAGTACTTCGGTACACAACTT 152
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
SNVLTR
LOCUS SNVLTR 180 bp DNA VRL 17-JUN-1992
DEFINITION spleen necrosis virus LTR.
ACCESSION X59450
VERSION X59450.1 GI:61992
KEYWORDS
SOURCE
    ORGANISM Spleen necrosis virus.
    Spleen necrosis virus
    Viruses; Retroid viruses; Retroviridae; Mammalian type C
    retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 180)
AUTHORS Iwasaki,K. and Temin,H.M.
TITLE The efficiency of RNA 3'-end formation is determined by the
distance between the cap site and the poly(A) site in spleen
necrosis virus
```


2000 AUG 8 10:15 AM

THIS PAGE BLANK (USPTO)

us-09-214-124-2.rn9

```

OM nuclelc - nuclelc search, using sw model
                        Search time 205.61 seconds
Run On: August 7, 2000, 09:21:12 ; (without alignments)
                        703.327 Million cell updates/sec

```

```

sequence:
scoring table:
IDENTITY-NUC      Gapext 1.0
Gapop 10.0 ,      residues
311585 seqs,      125096042
searched:          parameters:
623170

```

```

Total number of hits satisfied: 1
Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : N_Geneseq_36:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

SUMMARIES

Descriptor

Sequence derived from

Sequence user (FPV)
Sequence virus (FPV)
Sequence box (FPV)

FOUR	vector
Retrovirus	vector

Retrovirus vector

RETROVIRUS (FE)

Fowlpox syncytia
Chicken pox

Phaseolus lanceolatus a

Slap...
Heat-resistant

Heat-resistant
cattle growth

Sequence of genes
secreted

Human derived a Human

Human SHOX (shox) sequ

Human CSF3 prot

Human mcl-1 gene
mouse SRY-rela

Mouse Gene:
BSSL/CEU
PTHI

M. grisea
polymorphic re

polymorphic re

Uricase human t
Novel

Sequence of bone mineral density measurements

Musca p. falciparum

Continuation of
Human P110 cl

1

ALIGNMENTS

Human IL-1ra BAC c
Partial Factor VII
Factor IX/Factor V
Coprinus cinereus
Enterococcus faeca
Human IL-1ra BAC c
Human IL-1ra SCAREC
Arabidopsis
Kidney injury asso
Allergen Alt a 12
DNA encoding a kri
DNA encoding a kri
Human acylcoenzyme

NO 9429437

~~Proceedings~~ Page 1

RESULT	1
V63779	standard; DNA; 578 BP.

[illegible]

Query Match	Similarity	Prev. Similarity	Mismatches
Best Local Match	100.0%	0	0
Matches 578; Conservative			

[illegible]

RESULT 2
 ID V63778 standard; DNA; 940 BP.
 AC V63778;
 DT 17-MAR-1999 (first entry)
 DE Sequence derived from the 5' end of REV-A.
 KW Reticuloendotheliosis virus type A.
 KM type C retrovirus; internal ribosome entry site; IRES; encapsidation;
 OS retroviral vector; gene therapy; ss.
 PN FR2762615-A1.
 PD 30-OCT-1998.
 PF 28-APR-1997; FR-005203.
 PR 28-APR-1997; FR-005203.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PI Darlix JL, Gabus Darlix C, Lopez LM;
 DR WPI: 99-037487/04.
 PT Expression vectors containing IRES and/or encapsidation enhancer -
 PT derived from type C retrovirus other than FMLV and MoMLV
 PS Claim 5; Page 31-32; 43pp; French.
 CC The present sequence is derived from the 5' end of the genomic RNA of
 CC Reticuloendotheliosis virus type A (REV-A). The specification also
 CC describes nucleotide sequence derived from all or part of the 5' end
 CC of the genomic RNA of a type C retrovirus other than Friend murine
 CC leukaemia virus (FMLV) and Moloney murine leukaemia virus (MoMLV). The
 CC 5' derived sequences are used as an internal ribosome entry site (IRES)
 CC in a vector for permitting or enhancing the encapsidation of a retroviral
 CC vector. The vectors can be used for gene therapy, production of a recombinant
 CC recombinant polypeptides or production of transgenic animals.
 CC Sequence 940 BP; 190 A; 225 C; 260 G; 265 U;
 SO

Page 2

RESULT 3

Query Match	Score	Location
1127 G;	1127	line above.
1175 E	1175	

07T 025
|||
|||

Db	1651	-CTTTTCTTCTATATCCTCAGATTGGCAGTGAGAGAGATTGTTGTTGTTGTTGGCT	1709
OY	121	ggcuaacuggguggggugagggugcgcgacuaaaccguauuucgaaacaaauug	180
Db	1710	GGCTACTGGGTGGGTAGGGATCCGGACTGAATCCGTAGTATTTCGTACAAATTGG	1769
OY	181	gggcugcuccggggaauccucccaucgycagaaagugccuaucguuucgaaaccggc	240
Db	1770	GGGCTCGTCCGGGATTCCTCCCAATCGGCAGAGGTGCCCTACTGTCTTCGAAC	1829
OY	241	gccgguaaagaaacuaucuuuuguaaccucgcgaggguuugggagaaucgagugcg	300
Db	1830	GCCGGTAAGTAAGTACTGATTTTGGTACCTCCGAGGGTTGGGAGGTTCCGAGTGGCG	1889
OY	301	ggacgucgcccgggaagcuaaccuccucagcagggagcggccuacucgagcucug	360
Db	1890	GGACGCTGCCGGGAAGCTCCACCTCCGCTCAGCAGGGGACGCCCTGCTGAGCTCTGTG	1949
OY	361	guuacugauuuguuuggaaccgucuccaagacguguaauaauaagucuguguuugug	420
Db	1950	GTAATCGATTGTTGTTGAACCGTCTTAAGACGGTGATTAATAAGTCGTGTTGTGTG	2009
OY	421	uuuuguuuuaaccuuuguuuguuucguacuuugucagcagcggccucgaaauugugc	480
Db	2010	TTTGTTTGTTAACCTTGTGTTGTTGTTGTCACTTGTCGACAGCGCCCTGCGAAT	2069
OY	481	ccacaaccgcgcggcuugcgaauaauacuuggagagucuuuuugccuccagugucuccgu	540
Db	2070	CCACACCGCGCGGCTTGCGAATAATACTTTGGAGAGTCTTTGCCCTCCAGTGTCTCCGT	2129
OY	541	uuuuacucgucucucucuccucuccgcccgggauggg	578
Db	2130	TTGTAATCGTCTCTCTCCCTCTCCGCGGGGATGGG	2167

RESULT 4
 ID Q76041 standard; DNA; 3878 BP.
 AC Q76041;
 DT 20-JUL-1995 (first entry)
 DE Retrovirus-vector pPOL15-R1.
 KW Retrovirus; vector; pPOL15-R1;
 KM gene transfer; gene therapy; ss.
 OS Spleen necrosis virus.
 PN M09429417-2
 PD 22-DEC-1994
 PF 07-JUN-1994; U06415.
 PR 07-JUN-1993; US-073345.
 RA (UYNE-) UNIV NEW JERSEY.
 PI Dornburg RC;
 DR WPI; 95-036467/05.
 PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
 PT ability to produce progeny virus, in helper cell which can infect
 PT host cell and form provirus
 PT Example; Page 17; 28pp; English.
 CC The universal retroviral vector pPOL15-R1 (given in Q76041) was
 CC obtained by replacing the SV40 promoter and hpt gene of pPOL1-R1
 CC (Q76038) with the multiple cloning site of pBluescript II KS. The
 CC vectors allow cell-type specific gene expression and eliminate risks
 CC of downstream activation of cellular proto-oncogenes.
 SO Sequence 3878 BP; 922 A; 948 C. 000 0 1000

Query Match	72.28;	Score 417.4;	DB 1;	Length 3878;
Best Local Similarity	64.08;	Pred. No. 3.5e-126;		
Matches 378;	Conservative 160;	W		

[illegible]

Db	744	AACTTTTCTTCTGAATCCTCAGATTGGCAGTGAAGAGAGATTGTTCGTGCTGTG	803
QY	118	GCUGCCUACUGUGUGG-guaggguccggacugaaucgguaguauucguacaacau	176
Db	804	CCTGGCCTACTGGGTGGCGCAGGGATCCGGACTGAATCCGTACTTCGTACAACAT	863
QY	177	uuggggucuguccgggaa-uuccucccaucggcagaagugccuacuguuuc-----	228
Db	864	TTGGGGCTCGTCCGGGATACCCCTCCCATCGCAGAGGTGCCAACTGCTCTTCCA	923
QY	229	---ucgaacucggcgccgguag-uagaucuuuguuuuguaaccug-cgaggguuug	283
Db	924	TTCTGCAACTCCGGCGCCGGTGAGTTAAGTACTTGATTTGGTACTTCGCCGAGGGTTG	983
QY	284	ggaggaucgga---guggcggagcugccggagcucacacucuccagcaggggac	340
Db	984	GGAGGATCGGAGTGGTGGCGGAGCGCTGCGGGAAGCTCACACCTCCGCTCAGCAGGGGAC	1043
QY	341	gcccucaucgagcucuguguaucugauuguuuguaagcugucuccaagacgugauaa	400
Db	1044	GCCCTGACCTGAGCTCTGTGTATCTGATTTGTTGAGCCCTCCCTAAGACGGTGATA-	1102
QY	401	uaaagucguuguuuguguuuguuuuguuacuuuguuuguuugcuaacugucagag	460
Db	1103	-CTAAGTCGTGGCTGTGTGTGTGTTGTTGTTGCCCTGTGTTGTCGTTGTTCGACAG	1161
QY	461	CGCCCGCGaaauugugugucccaacacgcgcgcugcgaaauaacuuugagagucuu	520
Db	1162	CGCCTTGCGAATGGTGTATCCACACCGCGCGCTTGGGAATAATACCTTTGGAGAGCCTT	1221
QY	521	uugcucacagugucuuuguaucugucucucucucucucucgcgcgcg 571	
Db	1222	TTGCTTCCAGTGTCTTCGCTGTGACTGTCCTCCTCTCCCTCTCCGGCCG 1272	

RESULT 5
ID Q76040
AC Q76040; 076040 standard; DNA; 5519 BP.
DT 20-JUL-1995 (first entry)
DE Retrovirus vector pPOL11-R3.
KW Retrovirus; vector; pPOL11-R3; spleen necrosis virus; SNV;
KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW long terminal repeat; LTR; encapsidation; gene transfer;
KW gene therapy; ss.
OS Spleen necrosis virus.
PN W09429437-A.
PD 22-DEC-1994.
PF 07-JUN-1994; U06415.
PR 07-JUN-1993; US-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example: Page 16-17; 28bp; English.
CC New recombination-free, highly efficient retroviral vectors
CC pPOL11-R1 (given in Q76038), pPOL11-R2 (Q76039) and pPOL11-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
SQ Sequence 5519 BP; 1264 A; 1420 C; 1456 G; 1170 T

```
Query Match          70.9%; Score 410; DB 1; Length 5519;  
Best Local Similarity    62.98; Pred. No. 1.1e-123;  
Matches      381; Conservative   161; Mismatches   30; Indels    34; Gaps  
                        8;  
QY       1 ggggucgcgcguccuacaca---uuuguuugagcgcgcgcgcacgaauucgaucuguuaaa 57  
         |||||:|||||:||| :||:::||||| |||||||||:::|||||:::|||
```


[illegible][illegible]

CC New recombination-free, highly efficient retroviral vectors
CC pPol11-R1 (given in Q76038), pPol11-R2 (Q76039) and pPol11-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
CC Sequence 5528 BP; 1265 A; 1423 C; 1457 G; 1383 T;
SQ Sequence 5528 BP; 1265 A; 1423 C; 1457 G; 1383 T;

[illegible]

PI Boyle DB, Couparr BEH, Gould AR, Hertig C;
 MPI; 99-167428/14.
 PR Recombinant vaccine against fowlpox virus, used to prevent the virus
 PT from reticuloendotheliosis in chickens
 PT occurrence in chickens.
 PT Claim 18; Fig 5; 127PP; English.
 PS The invention relates to a recombinant reticuloendotheliosis virus (REV) is
 CC acid does not encode an active REV sequence. The recombinant is used to
 CC whose genome does not contain any REV sequence. The chickens older than
 CC used to produce a vaccine against fowlpox virus in chickens. The chickens older than
 CC prevent the occurrence of fowlpox virus. Alternatively, chickens older than
 CC vaccinated at older than 1 day of age. Alternatively, chickens older than
 CC 1 day can be vaccinated with the vaccine subsequent to vaccination with a
 CC 1 day can be vaccinated with the vaccine subsequent to vaccination with a
 CC FPM (FPV mild strain) at 1 day of age. Prior vaccine and field strains
 CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
 CC near full length provirus of REV, and can give rise to infectious REV
 CC when transfected into cell cultures and when chickens are infected. The
 CC present invention provides the means by which a FPV vaccine free from
 CC contamination by REV may be produced. Sequences X26110-115 represent
 CC long terminal repeat sequences. The new recombinant FPV nucleic acid
 CC molecule can comprise a FPV LTR sequence selected from the above.
 SQ Sequence 536 BP; 145 A; 115 C; 133 G; 143 T;

[illegible]

RESULT 9
 X26109 standard; DNA; 512 BP.
 ID X26109 standard; DNA; 512 BP.
 AC X26109; (first entry)
 DT 20-MAY-1999 (first entry)
 DE Chicken syncytial virus provirus REV LTR sequence.
 KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
 KM vaccine; chicken; LTR; long terminal repeat; ss.
 OS Avian reticuloendotheliosis virus.
 PN WO9907852-A1.
 PS 18-FEB-1999. AU0628.
 PD 07-AUG-1998; AU-008454.
 PE 08-AUG-1997; AU-008454.
 PR (CSIR) COMMONWEALTH SCI & IND RES ORG. C;
 PA Boyle DB, Coupar BEH, Gould AR, Hertig C;
 PI WPI; 99-167428/14.
 DR Recombinant vaccine against fowlpox virus - is free of contamination
 PT from reticuloendotheliosis virus, used to prevent the virus
 PT from reticuloendotheliosis virus (FPV) whose nucleic
 PT occurrence 4; Fig 5; 127pp; English.
 PT Example 4; Fig 5; 127pp; English.
 PS The invention relates to a recombinant fowlpox virus (REV) and
 CS acid does not encode an active reticuloendotheliosis virus. The recombinant FPV is
 CC whose genome does not contain any REV sequence. The vaccine is used to
 CC acid does not encode an active reticuloendotheliosis virus. The vaccine is used to
 CC used to produce a vaccine against fowlpox virus in chickens. The chickens older than
 CC used to produce a vaccine against fowlpox virus in chickens. The chickens older than
 CC prevent the occurrence of fowlpox virus. Alternatively, chickens older than
 CC prevent the occurrence of fowlpox virus. Alternatively, chickens older than
 CC vaccinated at older than 1 day of age. Prior vaccine and field strains
 CC 1 day can be vaccinated at 1 day of age. Some of the FPV strains carry a
 CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
 CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
 CC near full length provirus of REV, and can give rise to infectious REV
 CC when transfected into cell cultures and when chickens are infected. The
 CC present invention provides the means by which a FPV vaccine free from

QY 370 ugnuguaagracgcgucccaagaacgugauaaauaagaucguguuuguguuuguuugu 429
::: :||| | :||| ||| | :|| |:: : :: | : :: : |
Db 502 TGTTTTGGAGCTTCCTAAACTGTAGTAACCAAGTAATATTAGTAACCTTATTCG 561

QY 430 uaccuuguguu 441
: |:| :| :|
Db 562 TTTCCTCTGATT 573

RESULT 12
T62570/c
ID T62570 standard; DNA; 1276 BP.
AC T62570;
DT 27-OCT-1997 (first entry)
DE Heat-resistant barley beta-amylase gene promoter.
KW Barley; heat resistant; promoter; genetic engineering; transgenic;
KW plant; improved maturation; ds.
OS Hordeum vulgare.
PN WO9702353-A1.
PD 23-JAN-1997.
PF 05-JUL-1996; J01866.
PR 05-JUL-1995; JP-191028.
PA (SAPB) SAPPORO BREWERIES.
PI Ito K, Kihara M, Okada Y, Yoshigi N;
DR WPI; 97-108966/10.
PT Expression promoter for genes inserted into plant seeds - e.g.
PT heat-resistant beta-amylase gene inserted into barley seeds
PS Claim 5; Page 17-18; 33pp; Japanese.
CC T62570 is a heat resistant promoter derived from a barley beta-amylase
CC (EC 3.2.1.2) gene. The promoter is used for the production of
CC transgenic barley plants and seeds. Such plants and seeds have
CC improved maturation properties.
SQ Sequence 1276 BP; 393 A; 237 C; 234 G; 412 T;

[illegible]

RESULT 13
T62572/c
ID T62572 standard; DNA; 2142 BP.
AC T62572;
DT 27-OCT-1997 (first entry)
DE Heat-resistant barley beta-amylase gene promoter containing DNA.
KW Barley; heat resistant; promoter; genetic engineering; transgenic;
KM plant; improved maturation; ds.
OS Hordeum vulgare.
PN WC9702353-A1.
PD 23-JAN-1997.
PF 05-JUL-1996; J01866.
PR 05-JUL-1995; JP-191028.
PA (SAPB) SAPORO BREWERIES.
PI Ito K, Kihara M, Okada Y, Yoshiyagi N;
DR WPI: 97-108966/10.
PT Expression promoter for genes inserted into plant seeds - e.g.
PT heat-resistant beta-amylase gene inserted into barley seeds
PS Example 6; Page 20-21; 33pp; Japanese.
CC T62572 is a DNA molecule comprising a heat resistant promoter derived
CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences
CC used for integration of this promoter into a reporter plasmid. The
CC plasmid is used for the production of transgenic barley plants
CC and seeds. Such plants and seeds have improved maturation properties.
SQ Sequence 2142 BP; 651 A; 482 C; 395 G; 614 T.

Query Match	5.38;	Score 30.6;	DB 1;	Length 2142;
Best Local Similarity	27.18;	Pred. No. 3.3;		
Matches	23;	Conservative	28;	Mismatches 34;

QY 368 auuguuuugagccgucuccaagacggugauaaauaagucguuguuuuguuuuguuu 427
|::||::| ||:|||| |::| :||| |:| :||::|
Db 136 ATTATGTTCAACGCTGCCACTAGAGTTCAAGTCTCAGACITTGACATTGTGGTTCGA 77

QY 428 guaaccuuguuuguucgucacu 452
|::||:: :::||:|:
Db 76 GTTCCCTAATTATTTCATGCCCTT 52

	Location/Qualifiers
ID N60157 standard; DNA; 2205 BP.	
AC N60157;	
DT 04-AUG-1991 (first entry)	
DE Cattle growth hormone (somatotropin) gene.	
KW Somatotropin; growth hormone; tissue plasminogen-activator; ss.	
OS Bos taurus.	
FH Key	
FT exon	324..336
FT FT	/tag= a
FT intron	/number= 1
FT FT	337..584
FT FT	/tag= b
FT exon	/number= 1
FT FT	585..745
FT FT	/tag= c
FT intron	/number= 2
FT FT	746..971
FT FT	/tag= d
FT exon	/number= 2
FT FT	972..1088
FT FT	/tag= e
FT intron	/number= 3
FT FT	1089..1317
FT FT	/tag= f
FT exon	/number= 3
FT FT	1318..1479
FT FT	/tag= g
FT intron	/number= 4
FT FT	1480..1753
FT FT	/tag= h
FT exon	/number= 4
FT FT	1754..1954
FT FT	/tag= i
FT poly(a)_signal	/number= 5
FT FT	2032..2037
FT FT	/tag= j
PN EP-173552-A.	
PD 05-MAR-1986.	
PF 23-AUG-1985; 306017.	
PR 24-AUG-1984; US-644306.	
PR 06-NOV-1984; US-668764.	
PR 26-JUL-1985; US-758517.	
PA (UPJO) UPJOHN CO.	
PI Goodwin EC, Palermo DP, Post LE, Roltman FM, Thomsen DR;	
PI Woychik RP;	
DR MPI; 86-063386/10.	
DR N-PSDB; N60157.	
PT New recombinant DNA for polypeptide expression in eukaryotic cells - contg. poly:adenylation signal, coding gene, promoter and selection marker	
PS Disclosure; Page 5; 32pp; English.	
CC This cattle growth hormone sequence is part of a circular recombinant DNA which is especially useful for the high-yield expression of tissue plasminogen-activator.	
CC Sequence 2205 BP; 432 A; 616 C; 702 G; 452 T;	
SQ	

	Query Match	5.3%;	Score 30.6;	DB 1;	Length 2205;	
	Best Local Similarity	43.2%;	Pred. No. 3.4;	Mismatches 59;	Indels 0;	Gaps 0;
	Matches 54;	Conservative 12;				
OY	259	gaauuuuguaccucgcgagggguuuuggaggaucggagugcgaggacucgccgggaagcu	318			
		::: : :				
Db	740	GAGTTTGTCAGCTCCCCGAGCGATCGCTCCTAGGGGTGGGGAGGCAGGAAGGGGTGAATCC	799			
OY	319	ccaccuccgcucacgacgaggggagcgcgcccaucucagcucucuguaucucgaauuguungug	378			
Db	800	ACACCCCCTCCACACACAGTGGAGGAACA CTGAGAGTTCAGCCGATTATTTATCCAAGTAGG	859			
OY	379	accgu 383				
		:				
Db	860	GATGT 864				

RESULT	15
ID	N40081 standard; DNA; 2205 BP.
AC	N40081;
DT	04-FEB-1992 (first entry)
DE	Sequence of genomic bovine growth hormone (BGH).
KW	Growth promoter; hormone; lactation; ss.
OS	Bos taurus.
FH	Key
FT	tata_signal
FT	Location/Qualifiers
FT	235..243
FT	/tag= a
FT	324..336
FT	/tag= b
FT	/product= signal peptide
FT	584..651
FT	/tag= c
FT	/product= signal peptide
FT	652..745
FT	/tag= d
FT	/product= mature peptide
FT	973..1089
FT	/tag= e
FT	/product= mature peptide
FT	1319..1480
FT	/tag= f
FT	/product= mature peptide
FT	1689..1859
FT	/tag= g
FT	/product= mature peptide
FT	1938..1943
FT	/tag= h
FT	polya_signal
PN	EP-112012-A.
PD	27-JUN-1984.
PF	01-NOV-1983; 306655.
PR	08-NOV-1982; US-439683.
PA	(UYCA-) CASE WESTERN UNIV.
PI	Rottmann FM, Woychik RP, Nilson JH;
DR	WPI; 84-160023/26.
DR	P-PsDB; P40104.
PT	Prodn. of bovine growth hormone - by cultivation of host
PT	microorganisms obtd. by recombinant DNA methods
PS	Claire 1; Page 10-12; 15pp; English.
CC	Genomic clones p lambda GH2R-2 and -3 are new. They are deposited
CC	with E.coli HB101 as host as NRRL B-15154 and B-15155 resp. The BGH
CC	genomic clones in NRRL B-15154 and B-15155 transformants can be used
CC	to direct the synthesis of BGH in a eucaryotic host.
SQ	Sequence 2205 BP; 434 A; 702 G; 452 T;

Query Match	5.3%	Score 30.6;	DB 1;	Length 2205;
Best Local Similarity	43.2%;	Pred. No. 3.4;		
Matches	54;	Mismatches	59;	Indels 0;
		Conservative		Gaps 0;

[illegible]

Db	739	GAGTTTGTAAGCTCCCGAGGGATGCGTCTAGGGGTGGGAGGCACGAAAGGGTGAATCC	798
QY	319	ccaccuccgcgcagcaggggacgcccugaucugagcucucuguaucugauuguuugy	378
Db	799	ACACCCCCCTCCACACAGTGGAGGAACTGAGGAGTTACGCCGATTTTATCCAGTTAGC	858
QY	379	accgu	383
Db	859	GATGT	863

```
Search completed: August 7, 2000, 09:21:25
Job time: 458 sec
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 10:53:39 ; Search time 4994.97 Seconds
(without alignments)
652.183 Million cell updates/sec

Title: US-09-214-124-2_COPY_265_578
Perfect score: 314
Sequence: 1 gguaccucgcgagggguuuug.....ccucucgcgcccgggauggg 314

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 5187315402 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_om:
4: gb_om:
5: gb_om:
6: gb_ph:
7: gb_pl1:
8: gb_pl1:
9: gb_pl1:
10: gb_pl2:
11: gb_pl3:
12: gb_ro:
13: gb_sts:
14: gb_sy:
15: gb_un:
16: em_fun:
17: em_hum1:
18: em_hum2:
19: em_in:
20: em_om:
21: em_om:
22: em_om:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_sy:
29: em_un:
30: em_v1:
31: gb_htg1:
32: gb_htg2:
33: gb_in1:
34: gb_in2:
35: em_bal:
36: em_bal:
37: em_hum3:
38: em_hum4:
39: gb_pr4:
40: gb_htg3:
41: gb_htg4:
42: gb_htg5:
43: gb_htg6:

44: gb_htg7:
45: em_htg1:
46: em_htg2:
47: em_htg3:
48: em_hum5:
49: gb_pl3:
50: gb_pr5:
51: gb_htg8:
52: gb_htg9:
53: gb_htg10:
54: gb_htg11:
55: gb_htg12:
56: gb_htg13:
57: gb_htg14:
58: gb_in3:
59: gb_htg15:
60: gb_htg16:
61: gb_htg17:
62: em_htg4:
63: em_htg5:
64: em_htg6:
65: em_htg7:
66: em_hum6:
67: gb_htg18:
68: gb_htg19:
69: gb_htg20:
70: gb_htg21:
71: gb_htg22:
72: gb_htg23:
73: gb_htg24:
74: gb_htg25:
75: gb_htg26:
76: gb_htg27:
77: gb_htg28:
78: gb_htg29:
79: gb_htg30:
80: gb_htg31:
81: gb_v11:
82: gb_v12:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	314	100.0	578	5 A83549	A83549 Sequence 2
2	314	100.0	940	5 A83548	A83548 Sequence 1
3	304.4	96.9	4292	82 AF006065	AF006065 Fowlpox v
4	264.6	84.3	1530	81 RESNVX	V01200 Spleen necr
5	149.4	47.6	891	81 REXXX1	V01204 Spleen necr
6	44.8	14.3	290	13 AU025522	AU025522 Rattus no
7	37.6	12.0	378	13 AU026236	AU026236 Rattus no
8	37.6	12.0	7218	5 I66494	I66494 Sequence 14
9	37.2	11.8	113440	55 AC020935	AC020935 Homo sapi
10	37	11.8	70480	39 AC006973	AC006973 Homo sapi
11	37	11.8	173618	57 AC019045	AC019045 Homo sapi
12	36.2	11.5	2041	12 MUSENZE1A	M88481 Mouse enzym
13	36	11.5	16832	42 AC014255	AC014255 Drosophill
14	36	11.5	88866	41 AC010696	AC010696 Drosophill
15	36	11.5	114958	51 AC008356	AC008356 Drosophill
16	36	11.5	119182	51 AC008234	AC008234 Drosophill
17	36	11.5	224230	34 AE003682	AE003682 Drosophill
18	35.8	11.4	186518	60 AC016813	AC016813 Homo sapi
19	35.4	11.3	72394	75 AC040913	AC040913 Homo sapi
20	35.2	11.2	176043	77 AC027072	AC027072 Homo sapi
21	35.2	11.2	213721	10 HS172B20	AL022319 Human DNA
22	35	11.1	155278	10 HSDJ655C4	AL080286 Human DNA
23	34.8	11.1	168613	40 AL139318	AL139318 Homo sapi
24	34.6	11.0	167656	31 AP001645	AP001645 Homo sapi

Nature 285: 550-554
1980

PNA S 70 (4) 1230-
1284, 1982

REFERENCE
1 (bases 1 to 4292)
Hertig, C., Coupar, B.E., Gould, A.R. and Boyle, D.B.
Field and vaccine strains of fowlpox virus carry integrated
sequences from the avian retrovirus, reticuloendotheliosis virus
Virology 235 (2), 367-376 (1997)
JOURNAL
MEDLINE
97428585
2 (bases 1 to 4292)
Hertig, C.H., Coupar, B.E.H., Gould, A.R. and Boyle, D.B.
Direct Submission
Submitted (30-MAY-1997) Division of Animal Health, CSIRO, 5
Portarlington Road, Geelong, Victoria 3213, Australia
JOURNAL
TITLE
Location/Qualifiers
FEATURES
source

CDS

/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC58238.1"
/db_xref="GI:2393893"

/translation="MATVARMYKINTTGTSCVLSKSLIPDSYNEYNIDDLKIKE
FIEISMORCFSTKSVTSTVLTENRTNRSISTSHDKNEPYEENGIMNIECYFA
CLEGSCIVNVLGDRQIDNISSESGFLMDVNTDVIDTKYVGLTTRIKVDAHVFY
ONVIMPEKNLFSQTNPNFILDITVODRNVLITISKYIYVNCDDKYDIFELKYL
VDNCKLPMPLIPLSKYDFTTDLVSKSENVKTVLSKVHTSMKSYNNDSLPAVAVK
YGTVTI"
1812..3311
/gene="gag"
1812..3311
/gene="gag"
/codon_start=1
/product="gag"
/protein_id="AAC58239.1"
/db_xref="GI:2393894"

gene

/translation="MGQAGSKGLTLPLECLIKNFSDFKRAAGDYGEDVDSFTLRICE
LEWTFEGVWPKEGTLDPRVAANRVNIFGNPHGPDVITYITWIDITIERKYLKDC
GCKPTGSKVLASQKVNPRVLPSPAPESPPMRRAQFLDERPLSPAPAPPYPEV
PAIAEGEGEQPDSTVMAVSPHTRSGLEFGAQSPGMYPIRETGERDMGGRMRTYV
PFTTSDLYMENQNPFSQADPDEVISLESVFYTHQPTWDDCCQLRTLTTEERERY
RTERREVRNDQGVYDTEREIEAQFPAIRPDWDPTNGRNDLERYRQILLGLRAA
ARKPTNLKITEVROGADSPAYLERLYQAYRTWSPIDPRAPENQAAIVIOFVSQA
PDIRKIKIDGFOGKSSELVAIAQKVFQREDPAKATHELTOKMAKVIILAAQESRAE
RGSKTPGKGRPLGKNQACAYCKEKGHWKKNCKLVSGATPVLYEELQ"
3312..4292
/gene="pro/pol"
3312..4292
/gene="pro/pol"
/codon_start=1
/product="protease/polymerase"
/protein_id="AAC58240.1"
/db_xref="GI:2393895"

RESULT 4
RESNVX 1530 bp DNA VRL 01-OCT-1996
LOCUS Spleen necrosis virus sequence from the end of the U3 region to
DEFINITION 1.92Kbp from the 5' end.
ACCESSION V01200 J02387
VERSION V01200.1 GI:61757
KEYWORDS integration site; provirus; terminal repeat.
SOURCE Spleen necrosis virus.
ORGANISM
REFERENCE 1 (bases 1 to 4292)
AUTHORS Shimotohno, K., Mizutani, S. and Temin, H.M.
TITLE Sequence of retrovirus provirus resembles that of bacterial
JOURNAL Nature 285 (5766), 550-554 (1980)
MEDLINE 80254544
REFERENCE 2 (bases 1 to 1530)
AUTHORS O'Rear, J.J. and Temin, H.M.
TITLE Spontaneous changes in nucleotide sequence in proviruses of spleen
JOURNAL necrosis virus, an avian retrovirus
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 79 (4), 1230-1234 (1982)
COMMENT 82174569
The meaning of the substitution at 739-749 is not completely clear
in [2].

FEATURES

source Location/Qualifiers
1..1530

/organism="Spleen necrosis virus"
/db_xref="taxon:11836"
313

/note="T is missing in clone 13"
314

/note="G is missing in clone 13"
315

/note="G is missing in clone 13"
327

/note="T is missing in clone 13"
367

/note="G is A in clone 14-44"
367

/note="A may be G"
374

/note="C is T in clone 32"
374

/note="C is T in clone 13"
375

BASE COUNT 1204 a 973 c 1064 g 1051 t
ORIGIN

Query Match 96.9%; Score 304.4; DB 82; Length 4292;
Best Local Similarity 68.5%; Pred. No. 5.8e-82;
Matches 215; Conservative 93; Mismatches 6; Indels 0; Gaps 0;


```
variation 375 /note="T is C in clone 13"
variation 383 /note="A is T in clone 32"
variation 383 /note="A is T in clone 13"
variation 474 /note="G is A in clone 13"
CDS 592. .>1530 /note="coding frame (aaq?)"
```

```

/sequence="
/protein_id="CAA24513.1"
/db_xref="GI:61758"
/db_xref="SWISS-PROT:P03342"
/translation="MGQAGSKGLTPLECILKNFSDFKKRAGDYGEDYDSFALRLICE
LEWPTFGVGMEKEGTLDKFVVAARNVTEGNPGHPDOVIYITWTDTITERPKYLKSC
GCKPHRTSKYLASQKVNPRPVLPSAPESPRIIRAQGFLEDRPLSPADAPPPYEV
SAIVEDTREGQPDSTVMTSPPHRTSGLEGAQPSGMYPLRETGERMTGRPMRTYV
PTTSDLVNKKNNQNPSSFSQAPDVISLSEVYFTHQPTWDDCQQLRLITTEERER
VRTESRRREVANDCGVYTDEREIEAQFATRPDWGS"

```

```

variation      594 /note="G is A in clone 63"
variation      732 /note="G is C in clone 13"
variation      739 . 749 /note="substitute with 135-188 in clone 13"
variation      949
variation      1029 /note="A is G in clone 63"
variation      1085 /note="C is A in clone 32"
variation      1122 /note="A is G in clone 13"
variation      1192 /note="A is G in clone 3-73"
variation      1192

```

variation	1192	/note="additional C in clone 3-73"
BASE COUNT	344 a	389 c 432 g 365 t
ORIGIN		

Query Match	84.3%;	Score 264.6;	DB 81;	Length 1530;
Best Local Similarity	66.68;	Pred. No. 6.8e-70;		
Matches 211; Conservative	87;	Mismatches 14;	Indels 5;	Gaps 2.

OY	1	gguaccucgcgagggguuuuggaggaaucyga--gugccggacgcugcccggygaagcucca	57
Dd	282	GGTACCTCGCGAGGGTTTGGAGGATCGGAGTGTTGGCGGACGCTGC GGGAAGCTCCA	341
OY	58	ccuccgcucacgacgagggagacgcccucaucugacucucuguguaucugauuuuguagacc	117
Dd	342	CCTCCGCTCAGCAGGGGACGCCCTTGACCTGAGCTCTGTGTATCTGATTGTGTGAGCC	401
OY	118	guccccaaagcaguguaauaauaagucuguguuuuguguuuuguuuuaaccuuugunu	177
Dd	402	GTCCCTAAGACGGGTGATA--CTAAGTCGTGGCTTGTGTGTTTGTGTGCTTGTGTGTTT	459
OY	178	guucgucacuugucagacagcgcccucgaaauugugugccacacccgcgcgcuugcgaa	237
Dd	460	GTTCTGCTGTTTGTGCACACAGCGCCTTGCGAATTGGTGTACCACACACCGCGGCTTGCGAA	519
OY	238	uaauacuunuggagagucuuuuugccuccacagucuuuccguuuuguacuucgucuccuuccc	297
Dd	520	TAAATACTTTGGAGAGCCTTTTGCCCTCCAGTGTCTCCGTCTGTACTCGTCCCTCCTCCC	579
OY	298	ucuccgcgcgcggygauggg	314
Dd	580	TCTCCGCGCCGGGATGGG	596

RESULT	5
REXXX1	
LOCUS	REXXX1
DEFINITION	Spleen necrosis virus (5' end) integrated in chicken cell.
	891 bp DNA VRL 06-APR-1993

ACCESSION	V01204	J02388	M12296
VERSION	V01204.1	GI:61793	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
FEATURES			

Spleen necrosis virus.
Spleen necrosis virus
Viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 891)
Shimotohno,K., Mizutani,S. and Temin,H.M.
Sequence of retrovirus provirus resembles that of bacterial
transposable elements
Nature 285 (5766), 550-554 (1980)
80254544
Location/Qualifiers
1
891

```

source
1. .691
/organism="Spleen necrosis virus"
/db_xref="taxon:11836"

source
1. .25
/note="cellular DNA"
/organism="Spleen necrosis virus"
/db_xref="taxon:11836"

source
1. .891
/organism="Gallus sp."
/db_xref="taxon:9036"

source
26. .891
/organism="Gallus sp."

```

BASE COUNT	ORIGIN
194 a	215 c
	246 g
	236 t

Query Match	47.6%;	Score 149.4;	DB 81;	Length 891;
Best Local Similarity	64.0%;	Pred. No. 6.8e-35;		
Matches 126;	Conservative 55;	Mismatches 11;	Indels 5;	Gaps 2;

[illegible]

RESULT	6			
LOCUS	AU025522			
DEFINITION				
ACCESSION	AU025522	290 bp	DNA	STS
VERSION	AU025522			
KEYWORDS	AU025522.1	GI:4515445		
SOURCE	STS.			
ORGANISM	Rattus norvegicus	DNA, clone:OT29.32/731c09.		
	Rattus norvegicus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (sites)			
AUTHORS	Watanabe,T.K., Hishigaki,H., Kanemoto,N., Miyakita-Mizoguchi,A.			
	Oga,K., Okuno,S., Ono,T., Tsuji,A., Hayashi,H., Adachi,M.,			
	Yamasaki,Y., Iriye,Y., Takahashi,E., Takagi,T., Nakamura,Y. and			
	Tanigami,A.			
TITLE	The large-scale mapping of rat microsatellite markers			
JOURNAL	Unpublished (1998)			
REFERENCE	2 (bases 1 to 290)			
				02-MAR-1999

[illegible][illegible]

Query Match	12.0%	Score 37.6	DB 5	Length 7218
Best Local Similarity	4.5%	Pred. No. 0.71		
Matches 13	Conservative 156	Mismatches 117	Indels 0	Gap
QY	17	uagggaagcagcgaaguggcgagcagcgcgggaagcuccaccuccgcgcagcaggag	7	
Db	1031	TTCCGAGCTTGCTGCACAGCTCAGGAGGAGCTTCGATYYYYYYYYYYYYYYYY	1	
QY	77	gcccgauacugagcucuguguaucugauugugugagccgucuccaagcagguaua	1	
Db	1091	YYY	1	
QY	137	uauaagucgugguuguguguuuguuuacccuuguguuugucgacuuugcgacag	1	
Db	1151	YYY	1	
QY	197	cgcccgcgcaauugugugcccacacgcgcggcuugcgaaauaacuuugagagucuu	2	
Db	1211	YYY	12	
QY	257	uugccuccagugucuccgguuugacugucuccuccuccuccucc	302	
Db	1271	YYY	1316	
RESULT	9			
LOCUS	AC020935			
DEFINITION	AC020935 113440 bp DNA	HTG	09-FEB-2000	
ACCESSION	Homo sapiens chromosome 5 clone CTD-2280P20, WORKING DRAFT			
VERSION	AC020935			
KEYWORDS	AC020935.2 GI:6957623			
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 113440)			
	DOE Joint Genome Institute.			

TITLE	Sequencing of Human Chromosome 5
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 113440)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Feb 11, 2000 this sequence version replaced g1:6691295.

```

Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 113440)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, US
On Feb 11, 2000 this sequence version replaced gi:6691295.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 99051 bases at least Q40
Consensus quality: 106221 bases at least Q30
Consensus quality: 108903 bases at least Q20
Estimated insert size: 113440; sum-of-contigs estimation
Estimated insert size: 115000; pulse field gel estimation
Quality coverage: 6.27x in Q20 bases; pulse field gel estimation
Quality coverage: 6.35x in Q20 bases; sum-of-contigs estimation

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	de preserved.	1025:	contig of 1025 bp in length
*	1		gap of unknown length
*	1026	2323:	contig of 1298 bp in length
*			gap of unknown length
*	2324	3467:	contig of 1144 bp in length
*			gap of unknown length
*	3468	4870:	contig of 1403 bp in length
*			gap of unknown length
*	4871	6484:	contig of 1614 bp in length
*			gap of unknown length
*	6485	8070:	contig of 1586 bp in length
*			gap of unknown length
*	8071	10565:	contig of 2495 bp in length
*			gap of unknown length
*	10566	13656:	contig of 3091 bp in length
*			gap of unknown length
*	13657	19366:	contig of 5710 bp in length
*			gap of unknown length
*	19367	30006:	contig of 10640 bp in length
*			gap of unknown length
*	30007	40387:	contig of 10381 bp in length
*			gap of unknown length
*	40388	54734:	contig of 14347 bp in length
*			gap of unknown length
*	54735	71288:	contig of 16554 bp in length
*			gap of unknown length
*	71289	113440:	contig of 42152 bp in length..

FEATURES	Location/Qualifiers
source	1. .113440 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="5" /clone="CTD-2280P20"
BASE COUNT	31786 a 25192 c 24437 g 32015 t 10 others
ORIGIN	

Query Match	11.8%;	Score 37.2;	DB 55;	Length 113440;
Best Local Similarity	27.2%;	Pred. No. 1.1;		
Matches	31;	Conservative	35;	Mismatches 48;
				Indels 0;
				Gaps 0;

Oy 83 aucugagcucuguguaucugauuguugugaccgucuccaagacguguaauaanaag 142
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 35316 ATTGGGGTATTTTATTCTGATTTTCATTAGACATTTAATTGTGTGTGAGAGTGTTG 35375

[illegible]

RESULT	10
AC006973/c	
LOCUS	
DEFINITION	AC006973 70480 bp DNA
ACCESSION	Homo sapiens PAC clone RP5-909F12 from 7q34-q36, complete sequence
VERSION	AC006973
KEYWORDS	AC006973.2 GI:4699968
SOURCE	HTG.
	human.

ORGANISM	REFERENCE
Homo sapiens	1 (bases 1 to 70480)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
Sulston, J. E. and Waterston, R.	
Toward a complete human genome sequence	
Genome Res. 8 (11), 1097-1108 (1998)	

MEDLINE	99063752
REFERENCE	2 (bases 1 to 70480)
AUTHORS	Glossip,D., Bauer,C., Ames,M. and Haakenson,W.
TITLE	The sequence of Homo sapiens PAC clone RP5-909F12
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 70480)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (05-MAR-1999) Genome Sequencing Center, Washington State University, 4444 Forest Park Parkway, St. Louis, MO 63108

JOURNAL	Submitted to: School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA
REFERENCE	4 (bases 1 to 70480)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (28-APR-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	5 (bases 1 to 70480)

REFERENCE	5 (bases 1 to 70480)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Apr 28, 1999 this sequence version replaced gi:4337282.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@watsn.wustl.edu
----- Summary Statistics
Center project name: H_DJ0909F12

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send

Db 19737 ATGTGATGTTGTTCTGACTATG 19713

RESULT 11

AC019045

LOCUS

DEFINITION Homo sapiens chromosome 11, clone RP11-63G23, WORKING DRAFT

AC019045 173618 bp DNA HTG 13-MAR-2000

SEQUENCE, 18 unordered pieces.

AC019045 GI:7230851

AC019045.4

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 173618)

TITLE Waterston, R.H.

JOURNAL The sequence of Homo sapiens clone

REFERENCE Unpublished

TITLE 2 (bases 1 to 173618)

JOURNAL Waterston, R.H.

REFERENCE Direct Submission

AUTHORS Submitted (30-DEC-1999) Genome Sequencing Center, Washington

TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

JOURNAL

COMMENT On Mar 13, 2000 this sequence version replaced gi:7021777.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----

Center project name: H_NH0063G23

----- Summary Statistics -----

Sequencing vector: M13; 88%

Sequencing vector: Plasmid; 12%

Chemistry: Dye-primer ET; 88% of reads

Chemistry: Dye-terminator Big Dye; 12% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 163884 bases at least Q40

Consensus quality: 167019 bases at least Q30

Consensus quality: 168966 bases at least Q20

Insert size: 192000; agarose-fp

Insert size: 171918; sum-of-contigs

Quality coverage: 3.64 in Q20 bases; agarose-fp

Quality coverage: 4.09 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 18 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1455: contig of 1455 bp in length

1456 1555: gap of unknown length

1556 4255: contig of 2700 bp in length

4256 4355: gap of unknown length

4356 7302: contig of 2947 bp in length

7303 7402: gap of unknown length

7403 10141: contig of 2739 bp in length

10142 10241: gap of unknown length

10242 16585: contig of 6344 bp in length

16586 16685: gap of unknown length

16686 21874: contig of 5189 bp in length

21875 21974: gap of unknown length

21975 27791: contig of 5817 bp in length

27792 27891: gap of unknown length

27892 34069: contig of 6178 bp in length

34070 34169: gap of unknown length

34170 39365: contig of 5196 bp in length

39366 39465: gap of unknown length

39466 45870: contig of 6405 bp in length

45871 45970: gap of unknown length.

45971 54439: contig of 8469 bp in length

54439 54539: gap of unknown length

54540 63883: contig of 9344 bp in length

63883 63983: gap of unknown length

63984 73653: contig of 9670 bp in length

73654 73753: gap of unknown length

73754 83830: contig of 10077 bp in length

83831 83930: gap of unknown length

83931 99389: contig of 15459 bp in length

99390 99489: gap of unknown length

99490 113715: contig of 14226 bp in length

113716 113815: gap of unknown length

113816 139051: contig of 25236 bp in length

139052 139151: gap of unknown length

139152 173618: contig of 34467 bp in length.

FEATURES

source

1. .173618

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/clone="RP11-63G23"

BASE COUNT 57346 a 30229 c 30623 g 53714 t 1706 others

ORIGIN

Query Match 11.8%; Score 37; DB 57; Length 173618;

Best Local Similarity 26.6%; Pred. No. 1.3;

Matches 29; Conservative 35; Mismatches 45; Indels 0; Gaps 0;

QY 78 ccugaucugagcucugcugaucugaucugaucugaucugaucugaucugauc 137

Db 111333 CCATTAAACCCCTGTTGACCTAAGAGGGTGTGTTAATTCTATATATGAGAAT 111392

QY 138 auaagucguuguuuguuuguuuguuuguuuguuuguuuguuuguuugacac 186

Db 111393 TTTCACGATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGACAC 111441

RESULT 12

MUSENZEL1

LOCUS

DEFINITION Mouse enzyme E1 (E1) pseudogene, complete cds.

ACCESSION M88481.1

VERSION M88481.1 GI:193034

KEYWORDS enzyme E1; pseudogene.

SOURCE Mus musculus (strain 163H) male DNA.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 2041)

TITLE Mitchell, M.J., Woods, D.R., Tucker, P.K., Opp, J.S. and Bishop, C.E.

JOURNAL Homology of a candidate spermatogenic gene from the mouse Y chromosome to ubiquitin-activating enzyme E1

MEDLINE Nature 354, 483-486 (1991)

REFERENCE 2 (bases 1 to 2041)

AUTHORS Tucker, P.K., Phillips, K.S. and Lundrigan, B.

TITLE A mouse Y chromosome pseudogene is related to human ubiquitin activating enzyme E1

JOURNAL Mamm. Genome (1992) In press

FEATURES

source

1. .2041

/organism="Mus musculus"

/strain="163H"

/db_xref="taxon:10090"

/sex="male"

join(372. .443,680. .877,1438. .1512,1617. .1799)

gene

exon

372. .443

/gene="E1"

/note="putative"

*	12387	13210:	contig of 824 bp in length
*	13211	14064:	contig of 854 bp in length
*	14065	14927:	contig of 863 bp in length
*	14928	15773:	contig of 846 bp in length
*	15774	16609:	contig of 836 bp in length
*	16610	17440:	contig of 831 bp in length
*	17441	18413:	contig of 973 bp in length
*	18414	20167:	contig of 1754 bp in length
*	20168	21209:	contig of 1042 bp in length
*	21210	22037:	contig of 828 bp in length
*	22038	22814:	contig of 777 bp in length
*	22815	24092:	contig of 1278 bp in length
*	24093	25156:	contig of 1064 bp in length
*	25157	26406:	contig of 1250 bp in length
*	26407	27514:	contig of 1108 bp in length
*	27515	28459:	contig of 945 bp in length
*	28460	29241:	contig of 782 bp in length
*	29242	30331:	contig of 1090 bp in length
*	30332	31473:	contig of 1142 bp in length
*	31474	32401:	contig of 928 bp in length
*	32402	33576:	contig of 1175 bp in length
*	33577	34586:	contig of 1010 bp in length
*	34587	35888:	contig of 1302 bp in length
*	35889	36743:	contig of 855 bp in length
*	36744	37587:	contig of 844 bp in length
*	37588	39002:	contig of 1415 bp in length
*	39003	40108:	contig of 1106 bp in length
*	40109	42175:	contig of 2067 bp in length
*	42176	43203:	contig of 1028 bp in length
*	43204	44114:	contig of 911 bp in length
*	44115	45143:	contig of 1029 bp in length
*	45144	46081:	contig of 938 bp in length
*	46082	47525:	contig of 1444 bp in length
*	47526	49383:	contig of 1858 bp in length
*	49384	51442:	contig of 2059 bp in length
*	51443	52699:	contig of 1257 bp in length
*	52700	53837:	contig of 1138 bp in length
*	53838	55185:	contig of 1348 bp in length
*	55186	56606:	contig of 1421 bp in length
*	56607	57488:	contig of 882 bp in length
*	57489	59045:	contig of 1557 bp in length
*	59046	60348:	contig of 1303 bp in length
*	60349	62107:	contig of 1759 bp in length
*	62108	63171:	contig of 1064 bp in length
*	63172	64595:	contig of 1424 bp in length
*	64596	66301:	contig of 1706 bp in length
*	66302	67886:	contig of 1585 bp in length
*	67887	69777:	contig of 1891 bp in length
*	69778	71421:	contig of 1644 bp in length
*	71422	73685:	contig of 2264 bp in length
*	73686	75828:	contig of 2143 bp in length
*	75829	77946:	contig of 2118 bp in length
*	77947	80276:	contig of 2330 bp in length
*	80277	82214:	contig of 1938 bp in length
*	82215	83882:	contig of 1668 bp in length
*	83883	86060:	contig of 2178 bp in length
*	86061	88866:	contig of 2806 bp in length.

FEATURES	location/Qualifiers
source	1. .88866
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone="RPC198-154F9"
BASE COUNT	24367 a 20106 c 19195 g 25107 t 91 others
ORIGIN	

Query Match	11.5%;	Score 36;	DB 41;	Length 88866;
Best Local Similarity	25.0%;	Pred. No. 2.5;		
Matches	23;	Conservative	34;	Mismatches 35; Indels 0; Gaps 0;

OY 94 gugguaucugauuguugacgcgucccaagacgcyuganaaaagaucguguuugu 153
 | : : | : | : | : | : | : | : | : | : | : | :
Db 23758 GTTCCTTCTGCAAGTGTTGTCCCTGTGCCTCAGACTCGTGTGTGTGTTTTTTGGT 23817

QY 154 guguuguuuguaaccuuguuuguaucguca 185
|:::|:::|:::|:::|:::|
Db 23818 GTGTTTGTGGTTGTGTGTGTGTTCGTCA 23849

RESULT	15
AC008356/c	
LOCUS	AC008356 114958 bp DNA HTG 15-FEB-2000
DEFINITION	Drosophila melanogaster chromosome 3 clone BACR43K14 (D968) RPCI-98 85D-85D strain V; cn bw sp, *** SEQUENCING IN PROGRESS

ACCESSION	AC008356
VERSION	AC008356.3
KEYWORDS	HTG: HTGS_PHASE1.
SOURCE	fruit fly.
	Drosophila melanogaster

REFERENCE
AUTHORS
1 (bases 1 to 114958)
Celnikier, S.E., Agbayan, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Culhane, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Culhane, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,

Moshrefi, M., Nixon, K., Pacied, J.W., Falk, J., ...
Sethi, H., Swirskas, R.R., Man, K.H., Webster, D.,
Richards, S., Yee, M., Yu, C. and Rubin, G.M.

TITLE	JOURNAL	REFERENCE	AUTHORS
Sequencing of <i>Drosophila melanogaster</i>	unpublished	2 (bases 1 to 114958)	Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G., Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacle, J.M., Park, S., Pfeiffer, B., Poon, L., Sequelra, A., Sethi, H., Snir, E., Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rubin, G.M.

TITLE	Direct Submission
JOURNAL	Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT	On Feb 15, 2000 this sequence version replaced gi:6435852. Information about this sequence, including its location

For further information about this submission, please visit our sequence and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.

- * NOTE: This is a 'working draft' sequence. It currently consists of 88 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* * de preserved.
* *
* * 1
* * 804: contig of 804 bp in length
* * 885 884: gap of unknown length
* * 885 1469: contig of 585 bp in length
* * 1470 1549: gap of unknown length
* * 1550 2258: contig of 709 bp in length
* * 2259 2338: gap of unknown length
* * 2339 3477: contig of 1139 bp in length
* * 3478 3557: gap of unknown length
* * 3558 4327: contig of 770 bp in length
* * 4328 4407: gap of unknown length
* * 4408 5500: contig of 1093 bp in length
* * 5501 5580: gap of unknown length
* * 5581 6232: contig of 652 bp in length
* * 6233 6312: gap of unknown length
* * 6313 7459: contig of 1147 bp in length

```


*	7460	7539:	gap of unknown length
*	7540	8334:	contig of 795 bp in length
*	8335	8414:	gap of unknown length
*	8415	9165:	contig of 751 bp in length
*	9166	9245:	gap of unknown length
*	9246	9884:	contig of 639 bp in length
*	9885	9964:	gap of unknown length
*	9965	11112:	contig of 1148 bp in length
*	11113	11192:	gap of unknown length
*	11193	12395:	contig of 1203 bp in length
*	12396	12475:	gap of unknown length
*	12476	13549:	contig of 1074 bp in length
*	13550	13629:	gap of unknown length
*	13630	14949:	contig of 1320 bp in length
*	14950	15029:	gap of unknown length
*	15030	16127:	contig of 1098 bp in length
*	16128	16207:	gap of unknown length
*	16208	17350:	contig of 1143 bp in length
*	17351	17430:	gap of unknown length
*	17431	19081:	contig of 1651 bp in length
*	19082	19161:	gap of unknown length
*	19162	20834:	contig of 1673 bp in length
*	20835	20914:	gap of unknown length
*	20915	21898:	contig of 984 bp in length
*	21899	21978:	gap of unknown length
*	21979	22671:	contig of 693 bp in length
*	22672	22751:	gap of unknown length
*	22752	23764:	contig of 1013 bp in length
*	23765	23844:	gap of unknown length
*	23845	24973:	contig of unknown length
*	24974	25053:	gap of unknown length
*	25054	25780:	contig of 1129 bp in length
*	25781	25860:	contig of 727 bp in length
*	25861	27141:	gap of unknown length
*	27142	27221:	contig of 1281 bp in length
*	27222	28275:	contig of 1054 bp in length
*	28276	28355:	gap of unknown length
*	28356	30242:	contig of 1887 bp in length
*	30243	30322:	gap of unknown length
*	30323	31945:	contig of 1623 bp in length
*	31946	32025:	gap of unknown length
*	32026	32928:	contig of 903 bp in length
*	32929	33008:	gap of unknown length
*	33009	35085:	contig of 2077 bp in length
*	35086	35165:	gap of unknown length
*	35166	36633:	contig of 1468 bp in length
*	36634	36713:	gap of unknown length
*	36714	38597:	contig of 1884 bp in length
*	38598	38677:	gap of unknown length
*	38678	41039:	contig of 2362 bp in length
*	41040	41119:	gap of unknown length
*	41120	42801:	contig of 1682 bp in length
*	42802	42881:	gap of unknown length
*	42882	44451:	contig of 1570 bp in length
*	44452	44531:	gap of unknown length
*	44532	46258:	contig of 1727 bp in length
*	46259	46338:	gap of unknown length
*	46339	47721:	contig of 1383 bp in length
*	47722	47801:	gap of unknown length
*	47802	50220:	contig of 2419 bp in length
*	50221	50300:	gap of unknown length
*	50301	51504:	contig of 1204 bp in length
*	51505	51584:	gap of unknown length
*	51585	52896:	contig of 1312 bp in length
*	52897	52976:	gap of unknown length
*	52977	55035:	contig of 2059 bp in length
*	55036	55115:	gap of unknown length
*	55116	57120:	contig of 2005 bp in length
*	57121	57200:	gap of unknown length
*	57201	59206:	contig of 2006 bp in length
*	59207	59286:	gap of unknown length
*	59287	61708:	contig of 2422 bp in length
*	61709	61788:	gap of unknown length

61789	63710:	contig of 1922 bp in length
63711	63790:	gap of unknown length
63791	65200:	contig of 1410 bp in length
65201	65280:	gap of unknown length
65281	67585:	contig of 2305 bp in length
67586	67665:	gap of unknown length
67666	70390:	contig of 2725 bp in length
70391	70470:	gap of unknown length
70471	72950:	contig of 2480 bp in length
72951	73030:	gap of unknown length
73031	76068:	contig of 3038 bp in length
76069	76148:	gap of unknown length
76149	80201:	contig of 4053 bp in length
80202	80281:	gap of unknown length
80282	88995:	contig of 8714 bp in length
88996	88975:	gap of unknown length
88976	89766:	contig of 691 bp in length
89767	89846:	gap of unknown length
89847	90501:	contig of 655 bp in length
90502	90581:	gap of unknown length
90582	91197:	contig of 616 bp in length
91198	91277:	gap of unknown length
91278	91882:	contig of 605 bp in length
91883	91962:	gap of unknown length
91963	92559:	contig of 597 bp in length
92560	92639:	gap of unknown length
92640	93335:	contig of 696 bp in length
93336	93415:	gap of unknown length
93416	94074:	contig of 659 bp in length
94075	94154:	gap of unknown length
94155	94793:	contig of 639 bp in length
94794	94873:	gap of unknown length
94874	95491:	contig of 618 bp in length
95492	95571:	gap of unknown length
95572	96257:	contig of 686 bp in length
96258	96337:	gap of unknown length
96338	96916:	contig of 579 bp in length
96917	96996:	gap of unknown length
96997	97645:	contig of 649 bp in length
97646	97725:	gap of unknown length
97726	98402:	contig of 677 bp in length
98403	98482:	gap of unknown length
98483	99131:	contig of 649 bp in length
99132	99211:	gap of unknown length
99212	99806:	contig of 595 bp in length
99807	99886:	gap of unknown length
99887	100488:	contig of 602 bp in length
100489	100568:	gap of unknown length
100569	101206:	contig of 638 bp in length
101207	101286:	gap of unknown length
101287	101991:	contig of 705 bp in length
101992	102071:	gap of unknown length
102072	102681:	contig of 610 bp in length
102682	102761:	gap of unknown length
102762	103386:	contig of 625 bp in length
103387	103466:	gap of unknown length
103467	104106:	contig of 640 bp in length
104107	104186:	gap of unknown length
104187	104842:	contig of 656 bp in length
104843	104922:	gap of unknown length
104923	105606:	contig of 684 bp in length
105607	105686:	gap of unknown length

```

Query Match          11.5%; Score 36; DB 51; Length 114958;
Best Local Similarity 31.5%; Pred. No. 2.5;
Matches 29; Conservative 28; Mismatches 35; Indels 0; Gaps 0;

QY 101  cugaugugugugagaccgucuccaagaacgugugauaauaagucguguguuuguguuug 160
      |: : ::::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57736 CTTGTGTTTGTGGCCCGCACACAAAGTGTCATTAATTATGCTGTGTGGCTGTGTTG 57677
      :::::| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 161  uuuguuaccuuguguuuguuugucacuuuguc 192
      :::::| | | | | | | | | | | | | | | | | | | | | | | | | |

```


Tue Aug 8 08:33:20 2000

us-09-214-124-2_copy_265_578.rge

Page 12

Db 57676 TTTGTTTGCTGCGCTTTTGTTCGCTTGTAGTGG 57645

Search completed: August 7, 2000, 10:55:16
Job time: 6090 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:28:43 ; Search time 172.61 Seconds
(without alignments)
250.232 Million cell updates/sec

Title: US-09-214-124-2_COPY_265_578
Perfect score: 314
Sequence: 1 gguaccucgagggguuuug.....ccucucgagcgagggg 314

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5C_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/5D_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/6_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
7: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.6	12.0	7218	1 US-08-232-463-14	Sequence 14, Appl
2	30.6	9.7	1276	4 US-08-793-599-1	Sequence 1, Appl
3	30.6	9.7	2142	4 US-08-793-599-3	Sequence 3, Appl
4	28.8	9.2	3934	5 US-09-226-568-18	Sequence 18, Appl
5	28.8	9.2	3946	1 US-08-077-848A-1	Sequence 1, Appl
6	28.8	9.2	3946	5 US-09-211-640-1	Sequence 1, Appl
7	28.8	9.2	3946	6 PCT-US94-03547-1	Sequence 1, Appl
8	28.8	9.2	11531	1 US-08-068-945A-1	Sequence 1, Appl
9	28.8	9.2	11531	2 US-08-442-806-1	Sequence 1, Appl
10	28.2	9.0	221	1 US-08-222-177A-17	Sequence 17, Appl
11	28.2	9.0	278	1 US-08-222-177A-52	Sequence 52, Appl
12	28	8.9	3240	2 US-08-162-081B-34	Sequence 34, Appl
13	28	8.9	3240	3 US-08-780-872-34	Sequence 34, Appl
14	28	8.9	3412	2 US-08-162-081B-32	Sequence 32, Appl
15	28	8.9	3412	3 US-08-780-872-32	Sequence 32, Appl
16	27.6	8.8	8224	4 US-09-010-398-14	Sequence 14, Appl
17	27.2	8.7	2277	2 US-08-676-967-2	Sequence 2, Appl
18	27.2	8.7	2277	2 US-08-676-974-2	Sequence 2, Appl
19	27.2	8.7	2277	4 US-09-098-487-2	Sequence 2, Appl
20	27.2	8.7	2643	4 US-08-781-802-11	Sequence 11, Appl
21	27.2	8.7	3545	4 US-08-781-802-9	Sequence 9, Appl
22	27.2	8.7	4315	4 US-08-781-802-1	Sequence 1, Appl
23	27.2	8.7	7208	5 US-09-166-186-107	Sequence 107, App
24	27	8.6	2281	2 US-08-164-614A-1	Sequence 1, Appl
25	27	8.6	2281	4 US-08-456-489B-1	Sequence 1, Appl
26	27	8.6	2281	6 PCT-US93-01720-1	Sequence 1, Appl

27	27	8.6	2488	2 US-08-279-270A-2	Sequence 2, Appl
28	26.8	8.5	20235	2 US-07-642-734C-3	Sequence 3, Appl
29	26.2	8.3	1155	5 US-08-860-368B-19	Sequence 19, Appl
30	26.2	8.3	2381	3 US-08-484-993B-9	Sequence 9, Appl
31	26.2	8.3	2381	4 US-08-484-158B-9	Sequence 9, Appl
32	26.2	8.3	2381	4 US-08-484-596A-9	Sequence 9, Appl
33	26.2	8.3	2381	4 US-08-480-150A-9	Sequence 9, Appl
34	26.2	8.3	2381	5 US-08-458-731-9	Sequence 9, Appl
35	26.2	8.3	2381	5 US-08-149-223A-9	Sequence 9, Appl
36	26.2	8.3	3566	5 US-08-689-421-32	Sequence 32, Appl
37	26.2	8.3	5159	4 US-08-146-930-3	Sequence 3, Appl
38	26.2	8.3	5159	6 PCT-US93-03993-3	Sequence 3, Appl
39	26.2	8.3	5690	3 US-08-447-464-2	Sequence 2, Appl
40	26.2	8.3	5690	3 US-08-716-679-2	Sequence 2, Appl
41	26.2	8.3	13011	4 US-08-791-849A-14	Sequence 14, Appl
42	26	8.3	1003	1 US-07-800-364B-13	Sequence 13, Appl
43	26	8.3	1003	3 US-07-989-847-11	Sequence 11, Appl
44	26	8.3	1498	7 RE34606-4	Patent No. RE34,60
45	26	8.3	1500	2 US-08-322-965-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match	12.0%;	Score 37.6;	DB 1;	Length 7218;
Best Local Similarity	4.58;	Pred. No. 0.0073;		
Matches 13;	Conservative 156;	Mismatches 117;	Indels 0;	Gaps 0;

[illegible]

RESULT 2
US-08-793-599-1/c
: Sequence 1, Application US/08793599

```

? GENERAL INFORMATION:
? APPLICANT: OKADA, YUKIO
? APPLICANT: YOSHIGI, NAOHIRO
? APPLICANT: ITO, KAZUTOSHI
? APPLICANT: KIHARA, MAKOTO
? TITLE OF INVENTION: TISSUE-SPECIFIC PROMOTER
? FILE REFERENCE: 2589-0056-0 PCT
? CURRENT APPLICATION NUMBER: US/08/793,599
? EARLIER FILING DATE: 1997-03-04
? EARLIER APPLICATION NUMBER: PCT/JP96/01866
? EARLIER FILING DATE: 1996-07-05
? EARLIER APPLICATION NUMBER: JP 191028/1995
? EARLIER FILING DATE: 1995-07-05
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 1276
? TYPE: DNA
? ORGANISM: Barley
US-08-793-599-1

```

Query Match	9.7%	Score 30.6;	DB 4;	Length 1276;
Best Local Similarity	27.1%;	Pred. No. 0.67;		
Matches	23;	Mismatches	34;	Indels 0;
				Gaps 0;

Oy 104 auuguuugugacccgucuccaagacgcgugauaaunaagugcguuuuguguuuguu 163
|:::||:|||:||||| |::|||::|
Db 136 ATTAATGTTCAAACGTCTCCACTAGAGTTCAGTCTCAGACTTGACATTGTGTTGCCAA 77

Oy 164 guaacccuuguuugucguacuu 188
|::||:::|::|:
Db 76 GTTTCCTTAATTATTTCATGCCTT 52

RESULT 3
US-08-793-599-3/c
; Sequence 3, Application US/08793599
; Patent No. 5952489
; GENERAL INFORMATION:
; APPLICANT: OKADA, YUKIO
; APPLICANT: YOSHIGI, NAOHIRO
; APPLICANT: ITO, KAZUTOSHI

```

1  APPLICANT: KIHARA, MAKOTO
2  TITLE OF INVENTION: TISSUE-SPECIFIC PROMOTER
3  FILE REFERENCE: 2589-0056-0 PCT
4  CURRENT APPLICATION NUMBER: US/08/793, 599
5  CURRENT FILING DATE: 1997-03-04
6  EARLIER APPLICATION NUMBER: PCT/JP96/01866
7  EARLIER FILING DATE: 1996-07-05
8  EARLIER APPLICATION NUMBER: JP 191028/1995
9  EARLIER FILING DATE: 1995-07-05
10 NUMBER OF SEQ ID NOS: 3
11 SOFTWARE: PatentIn Ver. 2.0
12 SEQ ID NO 3
13 LENGTH: 2142
14 TYPE: DNA
15 ORGANISM: Barley
16
17 US-08-793-599-3

```

Query Match	9.7%	Score 30.6;	DB 4;	Length 2142;
Best Local Similarity	27.1%;	Pred. No. 0.84;		
Matches	23;	Mismatches	34;	Indels 0;
		Conservative		Gaps 0;

OY

104 aauguauguagccgucuccagaacguauaaanaagucuguuuguuguuuuu 163
|:::||||| | |: ||| :||::|
Dd 136 ATTATGTTCAACGTCTCACACTAGAGTTCAGTCTCAGACTTGACATTGGTTGC AA 77

QY 164 guuaccuuguuuugucgucacuu 188
|::|:: ::::|::|::
Db 76 GTTTCCTAATTATTTCATGCGCTT 52

RESULT 4

```

US-09-226-568-18/c
; Sequence 18, Application US/09226568
; Patent No. 6001992
; GENERAL INFORMATION:
; APPLICANT: Ackermann, Elizabeth J.
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcussen, Eric G.
; TITLE OF INVENTION: Antisense Modulation of No. 6001992el Anti-apoptotic
; TITLE OF INVENTION: bcl-2-Related Proteins
; FILE REFERENCE: ISPH-0337
; CURRENT APPLICATION NUMBER: US/09/226,568
; CURRENT FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3934
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
;
US-09-226-568-18

```

Query Match	9.2%;	Score 28.8;	DB 5;	Length 3934;
Best Local Similarity	56.2%;	Pred. No. 4.3;		
Matches 27; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0;

QY 261 cuccagugucuuccguuuuacucgucuccucuccucucgcgg 308
|:|||| |: ||| : ||:||:||:||: | ||: | | ||
Db 600 CTCACGCACTGCCGTACAATTCTCCTCTCCCTCTGCGGG 553

RESULT 5
 US-08-077-848A-1/c
 ; Sequence 1, Application US/08077848A
 ; Patent No. 5470955
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig, Ruth W.
 ; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
 ;

;; TITLE OF INVENTION: POLYPEPTIDE
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Spensley Horn Jubas & Lubitz
;; STREET: 1880 Century Park East, Suite 500
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: USA
;; ZIP: 90067
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/077,848A
;; FILING DATE: 16-JUN-1993
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Ph.D., Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: PD-2845
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 455-5100
;; TELEFAX: (619) 455-5110
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3946 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; IMMEDIATE SOURCE:
;; CLONE: mcl-1
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 61..1110
;; OTHER INFORMATION:
;; OTHER INFORMATION: /note= "When nucleotide 740 = C,
;; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
US-08-077-848A-1

Query Match 9.2%; Score 28.8; DB 1; Length 3946;
Best Local Similarity 56.2%; Pred. No. 4.3;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 261 cuccagugucuuuccguuuuacucucucucucucucucucgcgcgcg 308
Db 600 CTCACGCGACTGCCGTACACTCTCTCTCTCTCTCTCTCTCTCTCTGCGCG 553

RESULT 6
US-09-211-640-1/c
; Sequence 1, Application US/09211640
; Patent No. 6020466
; GENERAL INFORMATION:
; APPLICANT: Craig, Ruth W.
; TITLE OF INVENTION: ANTIBODIES WHICH SPECIFICALLY BIND mcl-1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/211,640
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/441,375
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Ph.D., Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: PD-2845
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 455-5100
;; TELEFAX: (619) 455-5110
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3946 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; IMMEDIATE SOURCE:
;; CLONE: mcl-1
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 61..1110
;; OTHER INFORMATION:
;; OTHER INFORMATION: /note= "When nucleotide 740 = C,
;; OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
US-09-211-640-1

Query Match 9.2%; Score 28.8; DB 5; Length 3946;
Best Local Similarity 56.2%; Pred. No. 4.3;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 261 cuccagugucuuuccguuuuacucucucucucucucucucgcgcgcg 308
Db 600 CTCACGCGACTGCCGTACACTCTCTCTCTCTCTCTCTCTCTCTCTGCGCG 553

RESULT 7
PCT-US94-03547-1/c
; Sequence 1, Application PC/TUS9403547
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: MYELOID CELL LEUKEMIA ASSOCIATED GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03547
; FILING DATE: 31-MAR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:


```

SEQUENCE CHARACTERISTICS:
LENGTH: 3946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mcl-1
FEATURE:
NAME/KEY: CDS
LOCATION: 61..1110
OTHER INFORMATION: /note= "when nucleotide 740 = C,
OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
OTHER INFORMATION: acid 227 = V."
PCT-US94-03547-1

```

Query Match	9.28;	Score 28.8;	DB 6;	Length 3946;
Best Local Similarity	56.28;	Pred. No. 4.3;		
Matches	27;	Mismatches	12;	Indels 0;
		Conservative		Gaps 0;

OY	261	cuccagugucucuccguuuuagucucucucuccucuccgagccgg	308
Db	600	CTCCAGCAGCTGCCGCTACAACTCGTCTCTCTCTCTCTGCTGCGG	553

RESULT 8
 US-08-068-945A-1
 ; Sequence 1, Application US/08068945A
 ; Patent No. 5616483
 ; GENERAL INFORMATION:
 ; APPLICANT: Bjursell, Gunnar
 ; APPLICANT: Carlsson, Peter
 ; APPLICANT: Enerback, Sven
 ; APPLICANT: Hansson, Jennart
 ; APPLICANT: Lidberg, Ulf
 ; APPLICANT: Nilsson, Jeanette
 ; APPLICANT: Tornell, Jan
 ; TITLE OF INVENTION: New DNA Sequences
 ; NUMBER OF SEQUENCES: 58
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: White & Case
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2787
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/068,945A
 ; FILING DATE: 27-MAY-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9201809-2
 ; FILING DATE: 11-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9201826-6
 ; FILING DATE: 12-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9202088-2
 ; FILING DATE: 03-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9300902-5
 ; FILING DATE: 19-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: sterner, Richard J.
 ; REGISTRATION NUMBER: 35,372
 ; REFERENCE/DOCKET NUMBER: 1103326-052
 ; TELECOMMUNICATION INFORMATION:
 ;

```

1 TELEPHONE: (212)819-8783
2 TELEFAX: (212)354-8113
3 INFORMATION FOR SEQ ID NO: 1:
4 SEQUENCE CHARACTERISTICS:
5 LENGTH: 11531 base pairs
6 TYPE: nucleic acid
7 STRANDEDNESS: double
8 TOPOLOGY: linear
9 MOLECULE TYPE: DNA (genomic)
10 ORIGINAL SOURCE:
11 ORGANISM: Homo sapiens
12 TISSUE TYPE: Mammary gland
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
16 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
17 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
18 FEATURE:
19 NAME/KEY: mat_peptide
20 LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
21 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
22 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
23 OTHER INFORMATION: /EC_number=3.1.1.1
24 OTHER INFORMATION: /product="Bile salt-Stimulated Lipase"
25 FEATURE:
26 NAME/KEY: 5'UTR
27 LOCATION: 1..1640
28 FEATURE:
29 NAME/KEY: TATA_signal
30 LOCATION: 1611..1617
31 FEATURE:
32 NAME/KEY: exon
33 LOCATION: 1641..1727
34 FEATURE:
35 NAME/KEY: exon
36 LOCATION: 4071..4221
37 FEATURE:
38 NAME/KEY: exon
39 LOCATION: 4307..4429
40 FEATURE:
41 NAME/KEY: exon
42 LOCATION: 4707..4904
43 FEATURE:
44 NAME/KEY: exon
45 LOCATION: 6193..6323
46 FEATURE:
47 NAME/KEY: exon
48 LOCATION: 6501..6608
49 FEATURE:
50 NAME/KEY: exon
51 LOCATION: 6751..6868
52 FEATURE:
53 NAME/KEY: exon
54 LOCATION: 8335..8521
55 FEATURE:
56 NAME/KEY: exon
57 LOCATION: 8719..8922
58 FEATURE:
59 NAME/KEY: exon
60 LOCATION: 10124..10321
61 FEATURE:
62 NAME/KEY: exon
63 LOCATION: 10650..11490
64 FEATURE:
65 NAME/KEY: 3'UTR
66 LOCATION: 11491..11531
67
68 US-08-068-945A-1

```

Query Match	9.28;	Score 28.8;	DB 1;	Length 11531;
Best Local Similarity	33.68;	Pred. No. 6.9;		
Matches	51;	Conservative	24;	Mismatches 77;
				Indels 0;
				Gaps 0;

QY 41 gcugccgggaagcuccaccgucagcagggagcccgucugucugagcucuguguan 100
DB 7107 GCTCACTGCACACCTCCACCTCTGCTTCAAGTGTCTCTGACTCAACCTCCCATGTAG 7166
QY 101 cugauuguuugagccgucuccaagacguguaauaauagucguguuuguuug 160
DB 7167 CTGGACTACAGCACATGCCACCATGATTAATTTCTGTGTCTTAGTAGGAT 7226
QY 161 uuuguuacccuguuuguuugucgucacugucg 192
DB 7227 GGAGTTTCATCGTGTAGCTAGGATGATCTCG 7258

RESULT 9

US-08-442-806-1
Sequence 1, Application US/08442806
Patent No. 5716817

GENERAL INFORMATION:

APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Enerback, Sven
APPLICANT: Hansson, Lennart
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
TITLE OF INVENTION: Genomic DNA Sequences
TITLE OF INVENTION: Encoding Human BSSL/CEL
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442, 806
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068, 945
FILING DATE: 27-MAY-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9202088-2
FILING DATE: 03-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9300902-5
FILING DATE: 19-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-052

TELEPHONE: (212)819-8783
TELEFAX: (212)354-8113

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:

ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:

NAME/KEY: CDS

LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
FEATURE:

NAME/KEY: mat_peptide

LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /EC_number=3.1.1.1
FEATURE: /product="Bile Salt-stimulated Lipase"

NAME/KEY: 5'UTR

LOCATION: 1..1640
FEATURE:

NAME/KEY: TATA_signal

LOCATION: 1611..1617
FEATURE:

NAME/KEY: exon

LOCATION: 1641..1727
FEATURE:

NAME/KEY: exon

LOCATION: 4071..4221
FEATURE:

NAME/KEY: exon

LOCATION: 4307..4429
FEATURE:

NAME/KEY: exon

LOCATION: 4707..4904
FEATURE:

NAME/KEY: exon

LOCATION: 6193..6323
FEATURE:

NAME/KEY: exon

LOCATION: 6501..6608
FEATURE:

NAME/KEY: exon

LOCATION: 6751..6868
FEATURE:

NAME/KEY: exon

LOCATION: 8335..8521
FEATURE:

NAME/KEY: exon

LOCATION: 8719..8922
FEATURE:

NAME/KEY: exon

LOCATION: 10124..10321
FEATURE:

NAME/KEY: exon

LOCATION: 10650..11490
FEATURE:

NAME/KEY: 3'UTR

LOCATION: 11491..11531
US-08-442-806-1

Query Match 9.2%; Score 28.8; DB 2; Length 11531;
Best Local Similarity 33.6%; Pred. No. 6.9;
Matches 51; Conservative 24; Mismatches 77; Indels 0; Gaps 0;

QY 41 gcugccgggaagcuccaccgucagcagggagcccgucugucugagcucuguguan 100
DB 7107 GCTCACTGCACACCTCCACCTCTGCTTCAAGTGTCTCTGACTCAACCTCCCATGTAG 7166
QY 101 cugauuguuugagccgucuccaagacguguaauaauagucguguuuguuug 160
DB 7167 CTGGACTACAGCACATGCCACCATGATTAATTTCTGTGTCTTAGTAGGAT 7226

Page 7

[illegible]

Tue Aug 8 08:33:21 2000

us-09-214-124-2_copy_265_578.mri

Page 9

```

1 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
2 COMPUTER: IBM PS/2
3 OPERATING SYSTEM: PC-DOS
4 SOFTWARE: Wordperfect
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/08/780,872
7 FILING DATE: 09-JAN-1997
8 CLASSIFICATION: 435
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: 08/162,081
11 FILING DATE: February 7, 1994
12 APPLICATION NUMBER: PCT/GB93/00761
13 FILING DATE: 13 April 1993
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Pasqualini, Patricia A.
16 REGISTRATION NUMBER: 34,894
17 REFERENCE/DOCKET NUMBER: LUD 5256
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: (212) 688-9200
20 TELEFAX: (212) 838-3884
21 INFORMATION FOR SEQ ID NO: 32:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 3412 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: single or double
26 TOPOLOGY: linear
27 FEATURE:
28 NAME/KEY: CDS
29 LOCATION: 1..3204
30 OTHER INFORMATION: /standard_name="CDS"
31 US-08-780-872-32

```

Query Match		8.9%	Score 28;	DB 3;	Length 3412;
Best Local Similarity		34.2%;	Pred. No. 7.4;		
Matches	26;	Conservative	20;	Mismatches	30; Indels 0; Gaps 0;
QY	140	aagucguguuuguuguuuacccuuguguuugucgaaccuugucagacgcgc	199		
		: : : : : : : : : :			
Db	2475	AATCAAGGCTTGATCTTCGAATGTACCTTATGGTTGTCATCGGTGACTGTGT	2534		
QY	200	ccugcgaaauugugugug	215		
		: : :			
Db	2535	GCGACTTATTGAGGTG	2550		

Search completed: August 7, 2000, 09:28:54
Job time: 908 sec

THIS PAGE BLANK (USPTO)

505 UACUUUGGAGAGUCUUUUGCCCUCCAGUGUCUUUCCGUUUUGUACUCGUCGCCUCCUCCCCUCU 564

PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example: Page 17; 28pp; English.
CC The universal retroviral vector pPOL15-R1 (given in Q76041) was
CC obtained by replacing the SV40 promoter and hpt gene of pPOL11-R1
CC (Q76038) with the multiple cloning site of pBluescript II KS. The
CC vectors allow cell-type specific gene expression and eliminate risks
CC of downstream activation of cellular proto-oncogenes.
SQ Sequence 3878 BP; 922 A; 948 C; 988 G; 1020 T.

Query Match	79.28;	Score 248.6;	DB 1;	Length 3878;
Best Local Similarity	65.88;	Pred. No. 1e-69;		
Matches 198;	Conservative 84;	Mismatches 114;		

QY	10	cgagggtuuugggagggaucgga--guaggcgagcugccggaagcucaccacucgcgcuc	66
Db	974	CGAGGGTTGGGAGGATCGGAGTGTGTGGCGGAGCGCTGCCGGGAAGCTCCACCTCCGCTC	1033
QY	67	agcaggggagccgccugaucugagcucuguguaucugaauguuugugagaccgucuccaag	126
Db	1034	AGCAGGGGACGCCCTGACCTGAGCTGTGTGTATCTGATGTGTGTGAGCCGTCCCTAAG	1093
QY	127	acgguugaaauaaagucuguguuuguguuuguuuacuuuguuuguuugucgucac	186
Db	1094	ACGGTGATA--CTAAGTCGTGGCTGTGTGTTGTTGTTGTCCTGTGTGTTGTTGTCGTGT	1151
QY	187	uuugcgcagcagcccccugcgaauuguguguccacacaccgcgcgcuugcgaauaaucuu	246
Db	1152	TTGTCGACAGCGCCTTGCGAATTGGGTACCCACACCGCGCGCTTGCGAATAATACTTT	1211
QY	247	ggagagucuuugcucuccagugucuuuccguuuuguaucuguaucucucucucucgcgc	306
Db	1212	GGAGAGCCTTTGGCCTCCAGTGTCTCCGCTGTACTGTCTCCTCCTCCTCCTCCGACC	1271
QY	307	g 307	
Db	1272	G 1272	

RESULT 5
Q76040
ID Q76040 standard; DNA; 5519 BP.
AC Q76040;
DT 20-JUL-1995 (first entry)
DE Retrovirus vector pPol11-R3.
KW Retrovirus; vector; pPol11-R3;
KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW long terminal repeat; LTR; encapsidation; gene transfer;
KW gene therapy; ss.
OS Spleen necrosis virus.
PN M09429437-A.
PD 22-DEC-1994.
PF 07-JUN-1994; U06415.
PR 07-JUN-1993; US-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example; Page 16-17; 28pp; English.
CC New recombination-free, highly efficient retroviral vectors
CC pPol11-R1 (given in Q76038), pPol11-R2 (Q76039) and pPol11-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.

5519 BP; 1264 A; 1420 C; 1456 G; 1379 T.

Query Match	76.0%	Score 238.6;	DB 1;	Length 5519;
Best Local Similarity	64.0%	Pred. No. 1.8e-66;		
Matches 208; Conservative	86.0%	Mismatches 14		

QY	1	g u a c c u c g c g a g g u u u g g a g a g a u c y g a --- g u g c g g a c g c u g c c g g a a g c u c c a	57
D b	956	G G T A C C T C G C G A G G G T T T G G A G A T C G G A G T G G T G G C G G A C G C T G C C G G A A G C T C A	1015
QY	58	c u c e c g c u c a g c a g g g a g a c g c c c u g a u c u a g c u c ----- u g u g n a u c u g a u	105
D b	1016	C C T C C G C T C A G C A G G G A C G C C C T G A C C T G A G C T C G A A T T C A G A T C T G T G T A T C T G A T	1075
QY	106	u g u u g u u g a c c g u c u c c e a a g a c g u g a u a a u a a g a u c g u g u u u g u u g u u u u g u	165
D b	1076	T G T T G T T G A G C C G T C C C T T A A G A C A C G G T G A T A - - C T A A G T C G T G G C T T G T G T G T T T G T T G T	1133
QY	166	u a c c u u g u u u u g u u u g u c a c a u u g u c g a c a g c g c c c u g c g a a u u g u g u c c a c a c a c g c	225
D b	1134	T G C C T T G T G T T T G T T G T C G T C G T T T G T G C A C A C G C C T T G C G A A T T G G T G T A C C A C A C A C C G C	1193
QY	226	g c g g c u u g c g a a u a u a c u u u g a g a g u c u u u u u g c u c c a g u g u c u u c c g u u u u g a c u g	285
D b	1194	G C G G C T T G C G A A T A T A C T T T G G A G A G C C T T T G C C T C A G T G T C T T C C G T C T G T A C T C G	1253
QY	286	u c u c u c u c u c c u c u c c g g c g g a 310	
D b	1254	T C C T C C T C C C T C T C C G G C C G G G A 1278	

```

RESULT      6
ID 076039 standard; DNA; 5525 BP.
AC 076039;
DT 20-JUL-1995 (first entry)
DE Retrovirus vector pPOL11-R2.
KW Retrovirus; vector; pPOL11-R2.
KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW long terminal repeat; LTR; encapsidation; gene transfer;
KW gene therapy; ss.
OS Spleen necrosis virus.
PN W09429437-A.
PD 22-DEC-1994.
PE 07-JUN-1994; U06415.
PR 07-JUN-1993; US-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example; Page 14-16; 28pp; English.
CC New recombination-free, highly efficient retroviral vectors
CC pPOL11-R1 (given in Q76038), pPOL11-R2 (Q76039) and pPOL11-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
SQ Sequence 5525 BP; 1264 A.; 1473 C.; 1457 G.

```

Query Match	76.0%;	Score 238.6;	DB 1;	Length 5525;
Best Local Similarity	64.0%;	Pred. No. 1.8e-66;		
Matches 208: Conservative	86.0%;	Mismatches 144;		

[illegible]


```
Db 1022 CCTCCGCTCAGACAGGGAGCGCCCTGACCTGAGCTCGAATTCAGATCTGTGTATCTGAT 1081
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 106 uguugugagaccgucuccaagaagcgugauaauaagucguguuuguguuuguuugu 165
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1082 TGTGTGAGCCGTCCTAGACGGTGATA--CTAAGTCGTGGCTGTGTGTGTGTGT 1139
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 uaccuuguguuugucacucugucacagcgcccgccgcaauugugugccacaccgc 225
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1140 TGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 226 gcgcuugcgaaauaauacuugagagagucuuugccuccagugucuccguuuuacucg 285
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1200 GCGGCTTGCGAATAATCTTTGGAGAGCCTTTTGCCTCCAGTGTCTCCGTCTGTACTCG 1259
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 uccuucucuccucgcccggcgga 310
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1260 TCCTCCTCTCCCTCCGCGCGGGA 1284
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 7
ID 076038 standard; DNA; 5528 BP.
AC 076038;
DE 20-JUL-1995 (first entry)
DE Retrovirus vector pPO111-R1.
KW Retrovirus; vector; pPO111-R1; spleen necrosis virus; SNV;
KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
KW long terminal repeat; LTR; encapsidation; gene transfer;
KW gene therapy; ss.
OS Spleen necrosis virus.
PN W09429437-A.
PD 22-DEC-1994.
PF 07-JUN-1994; U06415.
PR 07-JUN-1993; US-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example; Page 13-14; 28pp; English.
CC New recombination-free, highly efficient retroviral vectors
CC pPO111-R1 (given in Q76038), pPO111-R2 (Q76039) and pPO111-R3
CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
SQ Sequence 5528 BP; 1265 A; 1423 C; 1457 G; 1383 T;
```

```
Query Match 76.0%; Score 238.6; DB 1; Length 5528;
Best Local Similarity 64.0%; Pred. No. 1.8e-66;
Matches 208; Conservative 86; Mismatches 14; Indels 17; Gaps 3;
```

```
QY 1 gguaccucgagagguuuggaggaucgga---gugcgaggacgucgcccgggaagcucca 57
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 965 GGTACTCGGAGGGGTTTGGAGGATCGGAGTGTGGCGGGGACGCTGCCGGGAGCTCCA 1024
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 ccuccgucacgaggggagcgcccgucucugaguc-----uguguaucugau 105
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1025 CCTCCGCTCAGACAGGGAGCGCCCTGACCTGAGCTCGAATTCAATCTTGTGTATCTGAT 1084
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 106 uguugugagcgucuccaagaagcgugauaauaagucguguuuguguuuguuugu 165
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1085 TGTGTGAGCCGTCCTAAGACGGTGATA--CTAAGTCGTGGCTGTGTGTGTGTGTGT 1142
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 uaccuuguguuugucacucugucacagcgcccgcccgcaauugugugccacaccgc 225
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1143 TGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1202
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 226 gcgcuugcgaaauaauacuugagagagucuuugccuccagugucuccguuuuacucg 285
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 1203 GCGGCTTGCGAATAATCTTTGAGAGCCTTTTGCCTCCAGTGTCTCTGTACTCG 1262
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 uccuucucuccucgcccggcgga 310
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1263 TCCTCCTCTCCCTCCGCGCGGGA 1287
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 8
ID V74552 standard; DNA; 5000 BP.
AC V74552;
DE 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SEQ ID #241.
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT misc_feature 1021..1080
      /*tag= a
      /note= "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
```

```
FT misc_feature 2821..2880
      /*tag= b
      /note= "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
```

```
FT misc_feature 4621..4680
      /*tag= c
      /note= "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
```

```
PN EP-786519-A2.
PD 30-JUL-1997.
PF 07-JAN-1997; 100117.
PR 05-JAN-1996; US-009861.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
PI Rosen CA;
DR WPI; 97-374922/35.
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
PS Claim 1; Page 1025-1028; 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
SQ Sequence 5000 BP; 1377 A; 990 C; 641 G; 1808 T;
```

```
Query Match 10.2%; Score 32; DB 1; Length 5000;
Best Local Similarity 29.2%; Pred. No. 1;
```


FT	exon		4707.	.4904					
FT		/tag-	g						
FT	exon		6193.	.6323					
FT		/tag-	h						
FT	exon		6501.	.6608					
FT		/tag-	i						
FT	exon		6751.	.6868					
FT		/tag-	j						
FT	exon		8335.	.8521					
FT		/tag-	k						
FT	exon		8719.	.8922					
FT		/tag-	l						
FT	exon		10124.	.10321					
FT		/tag-	m						
FT	exon		10650.	.11490					
FT		/tag-	n						
FT			11491.	.11531					
FT		/tag-	o						
PN									
PD		WO9325669-A.							
PF		23-DEC-1993.							
PR		09-JUN-1993; SE0515.							
PR		11-JUN-1992; SE-001809.							
PR		12-JUN-1992; SE-001826.							
PR		03-JUL-1992; SE-002088.							
PA		19-MAR-1993; SE-000902.							
PI		(ASTR) ASTRA AB.							
PI		Bjursell KG, Carlsson PNI, Enerback CSM, Hansson SL;							
PI		Lidberg UFP, Nilsson JA, Tornell JBF, Toernell JBF;							
DR		Enerback CSM;							
DR		WPI; 94-007527/01.							
PT		P-PSDB; R45189.							
PT		DNA encoding bile salt-stimulated lipase/carboxyl ester lipase -							
PT		expressed by transgenic non-human mammals and used as human milk							
PS		substitute, contains intron sequences							
CC		Claim 1; Page 43-51; 76pp; English.							
CC		The primers (Q54223-28) are used to amplify fragments of the							
CC		BSSL / CEL gene, (Q54222). Which encodes a bile salt-stimulated							
CC		lipase / carboxyl ester lipase which when expressed by transgenic							
CC		animals can be used as a human milk substitute. The BSSL is designed							
CC		to pass through the stomach and is activated in the small intestine.							
CC		As it has a broad substrate specificity it can mediate the complete							
CC		digestion of most dietary lipids.							
SQ		Sequence 11531 BP; 2464 A; 3527 C; 3280 G; 2260 T;							

Query Match	Best Local Similarity	Score	DB	Length
Matches 51; Conservative 24; Mismatches 77; Indels 0; Gaps 0;	9.2%;	28.8;	1;	11531;
QY 41	gcugccggaagcuccaccuccgcgcucagggagcgcgccugaucugagcucuguguan	100		
Db 7107	gcctactgcAACCTCCACCTCCGTGGTTCAAGTAGATTCTGTACTCAACCTCCCATGTAG	7166		
QY 101	cugaaugugugugagaccgucuccaagaacguguaaauaauagucguguguguguguan	160		
Db 7167	CTGGGACTACAGGCACACATGCCACCATGCCACAGATAATTTTCTGTGTATTAGTAGGAT	7226		
QY 161	uuuguaaccuugugugugucgucacauugcg	192		
Db 7227	GGAGTTTCATCGTGTAGCTAGGATGATCTCG	7258		

RESULT	15
X32297	
ID	X32297 standard; DNA; 4795 BP.
AC	X32297;
DT	22-JUN-1999 (first entry)
DE	M. grisea PTH12 gene sequence.
KW	Fungal pathogenicity; imidazole glycerol phosphate dehydratase; PTH2;
KW	carbamoyl acetyl transferase; membrane associated pathogenicity protein;
KW	homeodomain transcription factor; PTH3; PTH11; PTH12; ss.
OS	Magnaporthe grisea.
PN	MO9913094-A2.

PD 18-MAR-1999.
PF 08-SEP-1998; U18730.
PR 10-SEP-1997; US-058460.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PI Sweigard JA;
DR WPI: 99-229247/19.
DR P-PSDB; Y06786.
PT New isolated fungal pathogenicity genes
PS Claim 33; Page 47-49; 50pp; English.
CC The invention relates to new isolated fungal pathogenicity genes,
CC designated PTH2, PTH3, PTH11 and PTH12 that are obtained from *Magnaporthe*
CC grisea. These novel genes encode proteins (Y06783-86) that are highly
CC homologous to the fungal carnitine acetyl transferase enzyme, fungal
CC imidazole glycerol phosphate dehydratase enzyme, fungal membrane
CC associated pathogenicity protein or fungal homeodomain transcription
CC factor respectively. The novel genes are implicated in fungal
CC pathogenicity. Inhibition of any of the genes PTH2, PTH3, PTH11 and PTH12
CC results in the reduction or elimination of the pathogenic phenotype of
CC the fungus. The isolated genes are useful in the design of screens to
CC identify inhibitors of the fungal pathogenic gene products. The present
CC sequence represents the PTH12 gene sequence. The present
SQ Sequence 4795 BP; 1220 A; 1338 C; 1067 G; 1170 T;

	Query Match	9.1%;	Score 28.6;	DB 1;	Length 4795;
	Best Local Similarity	32.8%;	Pred. No. 12;		
	Matches 22;	Conservative 21;	Mismatches 24;	Indels 0;	Gaps 0;
QY	121 uccaagacgugauaanaagaucguguguuuuguguguuuuguuacuuguguuuuguu	: :	:	: : :: : : : :	: : : :
Db	3828 TCAAAGCGGTGAATCCGCCAAGAAGGGTGTGTCGTAATTTTTTTTGTGTTGTGTAATG	:	:	:	:
QY	181 cgucaacu 187	: :			
Db	3888 TGCTCTCT 3894				

Search completed: August 7, 2000, 09:21:37
Job time: 470 sec

RESULT	15
X32297	
ID	X32297 standard; DNA; 4795 BP.
AC	X32297;
DT	22-JUN-1999 (first entry)
DE	M. grisea PTH12 gene sequence.
KW	Fungal pathogenicity; imidazole glycerol phosphate dehydratase; PTH2;
KW	carbamoyl acetyl transferase; membrane associated pathogenicity protein;
KW	homeodomain transcription factor; PTH3; PTH11; PTH12; ss.
OS	Magnaporthe grisea.
PN	MO9913094-A2.

Tue Aug 8 08:33:21 2000

us-09-214-124-2_copy_265_578.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

August 7, 2000, 11:56:12 ; Search time 3871.63 Seconds
(without alignments)

144.654 Million cell updates/sec

Title: US-09-214-124-2_COPY_452_578

Perfect score: 127

Sequence: 1 ugucgacagcgccucgcaa.....ccucucgcggcggauggg 127

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched:

5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters:

10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: gb_gss7:*
104: gb_gss8:*
105: gb_gss9:*
106: em_gss5:*
107: em_gss6:*
108: em_gss7:*
109: em_gss8:*
110: em_gss9:*
111: em_gss10:*
112: em_gss11:*
113: gb_gss10:*
114: gb_gss11:*
115: em_gss12:*
116: gb_gss12:*

SUMMARIES

ALIGNMENTS

SESSION	AI664738	GI:4775726
SESSION	AI664738	GI:4775726
WORDS	EST.	
WORDS	Trypanosoma cruzi.	
WORDS	Trypanosoma cruzi	
WORDS	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;	
ORGANISM	Trypanosoma; Schizotrypanum.	
REFERENCE	1 (bases 1 to 520)	
AUTHORS	Gonzalez Rey, E., Remisz, E., Delgado Garcia, A. and Gonzalez, A.	
TITLE	Characterization of ESTs from Trypanosoma cruzi epimastigotes	
JOURNAL	Unpublished (1998)	
COMMENT	Contact: Delgado Alberto Departamento de Biologia Molecular, Lab 303 Instituto de Parasitologia y Biomedicina Consejo Superior de Investigaciones Cientificas C/ Ventanilla No 11, E-18001, Granada, Spain Tel: 34 958 805058 Fax: 34 958 203323 Seq primer: T7 High quality sequence stop: 520.	
FEATURES	Location/Qualifiers	
source	1..520	
	/organism="Trypanosoma cruzi"	
	/strain="Cl - Brenner"	
	/db_xref="taxon:5693"	
	/clone="n114.r"	
	/clone_lib="T. cruzi epimastigote normalised cDNA library"	
	/cell_type="epimastigote"	
	/note="Site_1: EcorI; Site_2: NotI; cDNA library constructed with oligo dt primed epimastigote mRNA and cloned in pT7t318D phagemid with modified polylinker"	
BASE COUNT	143 a 89 c 181 g 107 t	
ORIGIN		
Query Match	26.1%; Score 33.2; DB 42; Length 520;	
Best Local Similarity	33.9%; Pred. No. 1.8;	
Matches	40; Conservative 25; Mismatches 53; Indels 0; Gaps 0;	
2Y	6 acagcgcccgcaauugugugccacacgcgcgcgcgcgaauaauacuuugagag 65	
Db	317 AAATCGCCACGCAATGTATTTGCCCTTAACCTGCCCCCTTCTCGTATTTTCCCG 258	
QY	66 uuuuuugccuacgugucucguuuugacucgucucucucucucucucgcgcgcgga 123	
Db	257 CCCTTCCTCTCCACTTCTTACCTTCCTCCTCCTCCTCCTCCTCATCGACCTGCA 200	
RESULT	2	
B19095	646 bp DNA	GSS 09-AUG-1997
LOCUS	B19095	T12K17-Sp6 TAMU Arabidopsis thaliana genomic clone T12K17,
DEFINITION	genomic survey sequence.	
ACCESSION	B19095	
VERSION	B19095.1	GI:2316999
KEYWORDS	GSS.	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta	
	Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;	
	Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 646)	
AUTHORS	Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.	
TITLE	BAC End Sequences at ATGC	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Ecker J.	


```

FEATURES
source
    LocalOut/Guanacaste
    1..643
    /organism="Lycopersicon esculentum"
    /cultivar="TA496"
    /db_xref="taxon:4081"
    /clone="CLEX2J20"
    /clone_lib="tomato root during/after fruit set, Cornell University"
    /tissue_type="root"
    /dev_stage="plants during and after fruit-set"
    /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XbaI; Site_3: PstI; Site_4: KpnI; Site_5: SalI; Site_6: BamHI; Site_7: SmaI; Site_8: NotI; Site_9: XhoI; Site_10: HindIII; Site_11: ClaI; Site_12: BclI; Site_13: PvuII; Site_14: MspI; Site_15: HpaI; Site_16: KpnI; Site_17: XbaI; Site_18: SmaI; Site_19: BclI; Site_20: PvuII; Site_21: MspI; Site_22: HpaI; Site_23: KpnI; Site_24: XbaI; Site_25: SmaI; Site_26: BclI; Site_27: PvuII; Site_28: MspI; Site_29: HpaI; Site_30: KpnI; Site_31: XbaI; Site_32: SmaI; Site_33: BclI; Site_34: PvuII; Site_35: MspI; Site_36: HpaI; Site_37: KpnI; Site_38: XbaI; Site_39: SmaI; Site_40: BclI; Site_41: PvuII; Site_42: MspI; Site_43: HpaI; Site_44: KpnI; Site_45: XbaI; Site_46: SmaI; Site_47: BclI; Site_48: PvuII; Site_49: MspI; Site_50: HpaI; Site_51: KpnI; Site_52: XbaI; Site_53: SmaI; Site_54: BclI; Site_55: PvuII; Site_56: MspI; Site_57: HpaI; Site_58: KpnI; Site_59: XbaI; Site_60: SmaI; Site_61: BclI; Site_62: PvuII; Site_63: MspI; Site_64: HpaI; Site_65: KpnI; Site_66: XbaI; Site_67: SmaI; Site_68: BclI; Site_69: PvuII; Site_70: MspI; Site_71: HpaI; Site_72: KpnI; Site_73: XbaI; Site_74: SmaI; Site_75: BclI; Site_76: PvuII; Site_77: MspI; Site_78: HpaI; Site_79: KpnI; Site_80: XbaI; Site_81: SmaI; Site_82: BclI; Site_83: PvuII; Site_84: MspI; Site_85: HpaI; Site_86: KpnI; Site_87: XbaI; Site_88: SmaI; Site_89: BclI; Site_90: PvuII; Site_91: MspI; Site_92: HpaI; Site_93: KpnI; Site_94: XbaI; Site_95: SmaI; Site_96: BclI; Site_97: PvuII; Site_98: MspI; Site_99: HpaI; Site_100: KpnI; Site_101: XbaI; Site_102: SmaI; Site_103: BclI; Site_104: PvuII; Site_105: MspI; Site_106: HpaI; Site_107: KpnI; Site_108: XbaI; Site_109: SmaI; Site_110: BclI; Site_111: PvuII; Site_112: MspI; Site_113: HpaI; Site_114: KpnI; Site_115: XbaI; Site_116: SmaI; Site_117: BclI; Site_118: PvuII; Site_119: MspI; Site_120: HpaI; Site_121: KpnI; Site_122: XbaI; Site_123: SmaI; Site_124: BclI; Site_125: PvuII; Site_126: MspI; Site_127: HpaI; Site_128: KpnI; Site_129: XbaI; Site_130: SmaI; Site_131: BclI; Site_132: PvuII; Site_133: MspI; Site_134: HpaI; Site_135: KpnI; Site_136: XbaI; Site_137: SmaI; Site_138: BclI; Site_139: PvuII; Site_140: MspI; Site_141: HpaI; Site_142: KpnI; Site_143: XbaI; Site_144: SmaI; Site_145: BclI; Site_146: PvuII; Site_147: MspI; Site_148: HpaI; Site_149: KpnI; Site_150: XbaI; Site_151: SmaI; Site_152: BclI; Site_153: PvuII; Site_154: MspI; Site_155: HpaI; Site_156: KpnI; Site_157: XbaI; Site_158: SmaI; Site_159: BclI; Site_160: PvuII; Site_161: MspI; Site_162: HpaI; Site_163: KpnI; Site_164: XbaI; Site_165: SmaI; Site_166: BclI; Site_167: PvuII; Site_168: MspI; Site_169: HpaI; Site_170: KpnI; Site_171: XbaI; Site_172: SmaI; Site_173: BclI; Site_174: PvuII; Site_175: MspI; Site_176: HpaI; Site_177: KpnI; Site_178: XbaI; Site_179: SmaI; Site_180: BclI; Site_181: PvuII; Site_182: MspI; Site_183: HpaI; Site_184: KpnI; Site_185: XbaI; Site_186: SmaI; Site_187: BclI; Site_188: PvuII; Site_189: MspI; Site_190: HpaI; Site_191: KpnI; Site_192: XbaI; Site_193: SmaI; Site_194: BclI; Site_195: PvuII; Site_196: MspI; Site_197: HpaI; Site_198: KpnI; Site_199: XbaI; Site_200: SmaI; Site_201: BclI; Site_202: PvuII; Site_203: MspI; Site_204: HpaI; Site_205: KpnI; Site_206: XbaI; Site_207: SmaI; Site_208: BclI; Site_209: PvuII; Site_210: MspI; Site_211: HpaI; Site_212: KpnI; Site_213: XbaI; Site_214: SmaI; Site_215: BclI; Site_216: PvuII; Site_217: MspI; Site_218: HpaI; Site_219: KpnI; Site_220: XbaI; Site_221: SmaI; Site_222: BclI; Site_223: PvuII; Site_224: MspI; Site_225: HpaI; Site_226: KpnI; Site_227: XbaI; Site_228: SmaI; Site_229: BclI; Site_230: PvuII; Site_231: MspI; Site_232: HpaI; Site_233: KpnI; Site_234: XbaI; Site_235: SmaI; Site_236: BclI; Site_237: PvuII; Site_238: MspI; Site_239: HpaI; Site_240: KpnI; Site_241: XbaI; Site_242: SmaI; Site_243: BclI; Site_244: PvuII; Site_245: MspI; Site_246: HpaI; Site_247: KpnI; Site_248: XbaI; Site_249: SmaI; Site_250: BclI; Site_251: PvuII; Site_252: MspI; Site_253: HpaI; Site_254: KpnI; Site_255: XbaI; Site_256: SmaI; Site_257: BclI; Site_258: PvuII; Site_259: MspI; Site_260: HpaI; Site_261: KpnI; Site_262: XbaI; Site_263: SmaI; Site_264: BclI; Site_265: PvuII; Site_266: MspI; Site_267: HpaI; Site_268: KpnI; Site_269: XbaI; Site_270: SmaI; Site_271: BclI; Site_272: PvuII; Site_273: MspI; Site_274: HpaI; Site_275: KpnI; Site_276: XbaI; Site_277: SmaI; Site_278: BclI; Site_279: PvuII; Site_280: MspI; Site_281: HpaI; Site_282: KpnI; Site_283: XbaI; Site_284: SmaI; Site_285: BclI; Site_286: PvuII; Site_287: MspI; Site_288: HpaI; Site_289: KpnI; Site_290: XbaI; Site_291: SmaI; Site_292: BclI; Site_293: PvuII; Site_294: MspI; Site_295: HpaI; Site_296: KpnI; Site_297: XbaI; Site_298: SmaI; Site_299: BclI; Site_300: PvuII; Site_301: MspI; Site_302: HpaI; Site_303: KpnI; Site_304: XbaI; Site_305: SmaI; Site_306: BclI; Site_307: PvuII; Site_308: MspI; Site_309: HpaI; Site_310: KpnI; Site_311: XbaI; Site_312: SmaI; Site_313: BclI; Site_314: PvuII; Site_315: MspI; Site_316: HpaI; Site_317: KpnI; Site_318: XbaI; Site_319: SmaI; Site_320: BclI; Site_321: PvuII; Site_322: MspI; Site_323: HpaI; Site_324: KpnI; Site_325: XbaI; Site_326: SmaI; Site_327: BclI; Site_328: PvuII; Site_329: MspI; Site_330: HpaI; Site_331: KpnI; Site_332: XbaI; Site_333: SmaI; Site_334: BclI; Site_335: PvuII; Site_336: MspI; Site_337: HpaI; Site_338: KpnI; Site_339: XbaI; Site_340: SmaI; Site_341: BclI; Site_342: PvuII; Site_343: MspI; Site_344: HpaI; Site_345: KpnI; Site_346: XbaI; Site_347: SmaI; Site_348: BclI; Site_349: PvuII; Site_350: MspI; Site_351: HpaI; Site_352: KpnI; Site_353: XbaI; Site_354: SmaI; Site_355: BclI; Site_356: PvuII; Site_357: MspI; Site_358: HpaI; Site_359: KpnI; Site_360: XbaI; Site_361: SmaI; Site_362: BclI; Site_363: PvuII; Site_364: MspI; Site_365: HpaI; Site_366: KpnI; Site_367: XbaI; Site_368: SmaI; Site_369: BclI; Site_370: PvuII; Site_371: MspI; Site_372: HpaI; Site_373: KpnI; Site_374: XbaI; Site_375: SmaI; Site_376: BclI; Site_377: PvuII; Site_378: MspI; Site_379: HpaI; Site_380: KpnI; Site_381: XbaI; Site_382: SmaI; Site_383: BclI; Site_384: PvuII; Site_385: MspI; Site_386: HpaI; Site_387: KpnI; Site_388: XbaI; Site_389: SmaI; Site_390: BclI; Site_391: PvuII; Site_392: MspI; Site_393: HpaI; Site_394: KpnI; Site_395: XbaI; Site_396: SmaI; Site_397: BclI; Site_398: PvuII; Site_399: MspI; Site_400: HpaI; Site_401: KpnI; Site_402: XbaI; Site_403: SmaI; Site_404: BclI; Site_405: PvuII; Site_406: MspI; Site_407: HpaI; Site_408: KpnI;
```

BASE COUNT	175 a	166 c	173 g	125 c
ORIGIN				

[illegible]

65 auctuauagcucaagugucuuaccguuagugacucugucucu 105

RESULT	11
AA321452/c	
LOCUS	
DEFINITION	
AA321452	246 bp
EST23928	Bone marrow
cell differentiation protein, mRNA sequence.	
EST 19-APR-1997	
similar to myeloid	

ACCESSION	AA321452	GI:1973779
VERSION	AA321452.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE	1 (bases 1 to 246)
AUTHORS	Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Weinstock, K.G., Gocayne, J.D.

Bult, C. J., Lee, N. N., Alexander, J. A., Brandon, R. C., Man-wai, C., White, O., Sutton, G., Blake, J. A., Fine, L. D., Clayton, R. A., Cline, T. R., Cotton, M. D., Earle-Hughes, J., Geoghagen, N. S., Fitzgerald, L. M., Fitchugh, W. M., Frichman, J. L., Hinkle, P. S., Jr., Glodex, A., Gnehm, C. L., Hanna, M. C., Hedblom, E., Marmaros, S. M., Merrick, J. M., Kelley, J. M., Kelley, J. C., Liu, L.-I., Nguyen, D. T., Pelligrino, S.

Phillips, C.A., Ryder, S.E., Scott, J.L., Saudak, D.W., Shattuck, M.L., Small, K.V., Spriggs, T.A., Utterback, T.R., Weldman, J.F., Wiley, Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Olsen, H., P. S. Weissner, J. Olsen, H.

TITLE Fraser, C.M. and Venter, J.C.
Initial assessment of human gene diversity and expression pattern
by analysis of cDNA sequence

JOURNAL
MEDLINE
COMMENT
Nature 377 (654/ Suppl.), 1997 (1997)
12140200
Other-ESTs: THC171424
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA

hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fátima Bonaldo. " 1 others

BASE COUNT	50	85	50	2	1	1
ORIGIN	a	c	g	d	e	f

Query Match	22.7%;	Score 28.8;	DB 86;	Length 276;
Best Local Similarity	56.2%;	Pred. No. 38;		
Matches	27;	Mismatches	12;	Gaps 0;
		Indels	0;	

QY	74	cuccagugucuccguuuuacucgucucucuccucucgagccgg	121
		: : : : : :	
Db	242	CTCCAGGACTGCCGGTACAACTCGTCTCTCTCTGCTGCGCG	195

RESULT	14
AA291667/c	
LOCUS	AA291667
DEFINITION	285 bp mRNA EST 16-MAY-1997
	z37c12.s1 Soares ovary tumor NbHOT Homo sapiens CDNA clone
	z37c12.s1 similar to SW:MCL1 HUMAN Q07820 INDUCED MYELOID
	z37c12.s1

ACCESSION AA291667
VERSION AA291667.1 GI:1939645

SOURCE
ORGANISM
human.
Homo sapiens
Chordata: Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

1 (bases 1 to 285)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Lennon, G., Marra, M., Martin, J.,

TITLE	JOURNAL	COMMENT
Washu-Merck EST Project 1997	Unpublished (1997)	Contact: Wilson RK
		Contact: [University School] of Medicine

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
possible reversed clone: similarity on wrong strand
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 267.

FEATURES	Location/Qualifiers
source	1. . 285

```

/clone_image="747054"
/clone_lib="Soares ovary tumor NbHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5
TGTTCACCATCTGAAGTGGGAGCGCCGCGGTTCCTTTTTCCTTTTTC
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

```

BASE COUNT	52 a	91 c	102 g	40 t
ORIGIN				

Query Match	22.7%	Score 28.8;	DB 23;	Length 285;
Best Local Similarity	56.28;	Pred. No. 38;		
Matches	27;	Mismatches	12;	Gaps 0;

QY	74	cuccagugucuccguuuuacucgucucucuccucucgcgcg	121
		: : : :	
Db	243	CTCCAGCGACTGCCGGTACAACTGTCCTCTCTGCTGGCGG	196

RESULT	15
AA883427/c	
LOCUS	317 bp mRNA
DEFINITION	EST
AA883427	27-MAR-1998
am25a12.s1	Soares-NFL_T-GBC_S1 Homo sapiens CDNA clone
accnt.1467938.3/	similar to SW:MCL1_HUMAN Q07820 INDUCED MYELOID

AA883427	AA883427.1	GI:2992957
ACCESSION		
VERSION		

SOURCE	ORGANISM	CHARACTER
	Homo sapiens	
	Chordata: Craniata: Vertebrata; Euteleostomi;	

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 317)	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap .
	National Cancer Institute,	Cancer Genome Anatomy Project (CGAP),
	NIH, Bethesda, Maryland	

JOURNAL Unpublished (1997)
COMMENT On Jan 5, 1998 this sequence replaced gi:2581350.
Contact: Robert Strausberg, Ph.D.

Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 205.
 Location/Qualifiers

FEATURES	Location/Qualifiers
source	1. .317
	1. .317 "Homo sapiens"

```

/clone1= "na2.1.1"
/clone1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19w, testis NHT, and B-cell
NCI-CGAP-GC81) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

```

BASE COUNT ORIGIN	63 a	98 c	115 g	120 e
1	100	100	100	100
2	100	100	100	100
3	100	100	100	100
4	100	100	100	100
5	100	100	100	100
6	100	100	100	100
7	100	100	100	100
8	100	100	100	100
9	100	100	100	100
10	100	100	100	100
11	100	100	100	100
12	100	100	100	100
13	100	100	100	100
14	100	100	100	100
15	100	100	100	100
16	100	100	100	100
17	100	100	100	100
18	100	100	100	100
19	100	100	100	100
20	100	100	100	100
21	100	100	100	100
22	100	100	100	100
23	100	100	100	100
24	100	100	100	100
25	100	100	100	100
26	100	100	100	100
27	100	100	100	100
28	100	100	100	100
29	100	100	100	100
30	100	100	100	100
31	100	100	100	100
32	100	100	100	100
33	100	100	100	100
34	100	100	100	100
35	100	100	100	100
36	100	100	100	100
37	100	100	100	100
38	100	100	100	100
39	100	100	100	100
40	100	100	100	100
41	100	100	100	100
42	100	100	100	100
43	100	100	100	100
44	100	100	100	100
45	100	100	100	100
46	100	100	100	100
47	100	100	100	100
48	100	100	100	100
49	100	100	100	100
50	100	100	100	100
51	100	100	100	100
52	100	100	100	100
53	100	100	100	100
54	100	100	100	100
55	100	100	100	100
56	100	100	100	100
57	100	100	100	100
58	100	100	100	100
59	100	100	100	100
60	100	100	100	100
61	100	100	100	100
62	100	100	100	100
63	100	100	100	100
64	100	100	100	100
65	100	100	100	100
66	100	100	100	100
67	100	100	100	100
68	100	100	100	100
69	100	100	100	100
70	100	100	100	100
71	100	100	100	100
72	100	100	100	100
73	100	100	100	100
74	100	100	100	100
75	100	100	100	100
76	100	100	100	100
77	100	100	100	100
78	100	100	100	100
79	100	100	100	100

Query Match	22.7%;	Score 28.8;	DB 32;	Length 317;
Best Local Similarity	56.2%;	Pred. No. 39;		
Matches	27;	Mismatches	9;	Indels 0; Gaps 0;

QY 74 cuccagugucuccgguuquacucgucucccuccucgcgg 121
| : | | | | : | : | : | : | : | : | : | : | : |
Db 239 CTCACGCGACTGCCGTACAACCTGTCCTCCTCCTCCTGCTGGCGG 192

Search completed: August 7, 2000, 11:56:19
Job time: 9753 sec

Tue Aug 8 08:33:23 2000

us-09-214-124-2_copy_452_578.rst

THIS PAGE BLANK (USPTO)


```

ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: PD-2845
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mcl-1
FEATURE:
NAME/KEY: CDS
LOCATION: 61..1110 /note= "when nucleotide 740 = C'
OTHER INFORMATION: amino acid 227 = A; when nucleotide 740 = T, amino
OTHER INFORMATION: acid 227 = V."
US-09-211-640-1
Query Match      22.7%; Score 28.8; DB 5; Length 3946;
Best Local Similarity 56.2%; Pred. No. 0.4;
Matches 27; Conservative 9; Mismatches 12; Indels 0; Gaps

```

RESULT 4
PCT-US94-03547-1/c
Sequence 1, Application PC/TUS9403547
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: MYELOID CELL LEUKEMIA ASSOCIATED GENE
TITLE OF INVENTION: MCL-1
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03547
FILING DATE: 31-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-2845
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:

QY 100 ccuccucuccucucgcgc 118
|||:|:||||
Db 1670 CCACTTGCCCTAATCCTGC 1688

RESULT 9

```

US-08-480-150A-9
; Sequence 9, Application US/08480150A
; Patent No. 5989550
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immuncontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,150A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/149,223
; FILING DATE: 09-NOV-1993
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 206..2353
US-08-480-150A-9

```

Query Match	20.6%;	Score 26.2;	DB 4;	Length 2381;
Best Local Similarity	39.2%;	Pred. No. 2.8;		
Matches	31;	Conservative	15;	Mismatches 33;
			Indels	0;
			Gaps	0;

[illegible]

RESULT 10

```

US-08-458-731-9
; Sequence 9, Application US/08458731
; Patent No. 6001599
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immuncontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,731
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELE: 25-3856
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 206..2353
; US-08-458-731-9

```

Query Match	20.6%	Score 26.2;	DB 5;	Length 2381;
Best Local Similarity	39.2%;	Pred. No. 2.8;		
Matches 31;	Conservative 15;	Mismatches 33;	Indels 0;	Gaps 0;

	Matches	31;	Conservative	15;	Mismatches	33;	Indels	0;	Gaps	0;
QY	40	cggcucugcg	aaauaacu	uuggagag	uucuuugc	uccaguguc	uuccguuug	uacucgu	99	
		:	:	:	:	:	:	:		
Db	1610	CTGCTGATAA	TATACCAATG	TCCAAAGTCT	TCTCCTCCCG	TGGGCTCAGT	GAGGCGCTGT	GT	1669	
QY	100	cuuccucuc	uccucucuc	gcgcgcgc	gcgcgcgc	gcgcgcgc	gcgcgcgc	gcgcgcgc	118	
		:	:	:	:	:	:	:		
Db	1670	CCACTTGCCCT	TATATCCTGC						1688	

RESULT 12

```

RESULT 12
US-08-480-552-18/c
; Sequence 18, Application US/08480552
; Patent No. 566550
;
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B.
; TITLE OF INVENTION: Genes And Genetic Elements Associated
; TITLE OF INVENTION: With Sensitivity To Chemotherapeutic Drugs
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 75 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,552
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,086
; FILING DATE: 09 MAR 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Keown, Wayne A.
; REGISTRATION NUMBER: 33,923
; REFERENCE/DOCKET NUMBER: 93,354
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/345-9100
; TELEFAX: 617/345-9111
;
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
;
US-08-480-552-18

```

RESULT 13

Query Match	20.5%;	Score 26;	DB 1;	Length 2389;
Best Local Similarity	36.5%;	Pred. No. 3.3;		
Matches 27; Conservative	17;	Mismatches 30;	Indels 0;	Gaps 0.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 11:56:08 ; Search time 3871.63 Seconds
(without alignments)

357.649 Million cell updates/sec

Title: US-09-214-124-2_COPY_265_578
Perfect score: 314

Sequence: 1 gguacucgcgagggguuuug.....ccucucgcggcggauggg 314

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 segs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*

44:	gb_est25:*
45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	gb_est30:*
50:	gb_est31:*
51:	gb_est32:*
52:	em_est20:*
53:	em_est21:*
54:	em_est22:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_est45:*
80:	gb_est46:*
81:	gb_est47:*
82:	em_est35:*
83:	em_est36:*
84:	em_est37:*
85:	gb_est48:*
86:	gb_est49:*
87:	gb_est50:*
88:	gb_est51:*
89:	gb_est52:*
90:	gb_est53:*
91:	gb_est54:*
92:	gb_est55:*
93:	gb_gss1:*
94:	gb_gss2:*
95:	gb_gss3:*
96:	gb_gss4:*
97:	em_gss1:*
98:	em_gss2:*
99:	em_gss3:*
100:	em_gss4:*
101:	gb_gss5:*
102:	gb_gss6:*
103:	gb_gss7:*
104:	gb_gss8:*
105:	gb_gss9:*
106:	em_gss5:*
107:	em_gss6:*
108:	em_gss7:*
109:	em_gss8:*
110:	em_gss9:*
111:	em_gss10:*
112:	em_gss11:*
113:	gb_gss10:*
114:	gb_gss11:*
115:	em_gss12:*
116:	gb_gss12:*


```

117:    gb_gss13: *
118:    gb_gss14: *
119:    gb_gss15: *
120:    gb_gss16: *
121:    gb_gss17: *
122:    gb_gss18: *
123:    gb_gss19: *
124:    em_gss13: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	40.8	13.0	807	122	CNS0119V	AL099997 Drosophill
C 2	39.4	12.5	1101	123	CNS017GS	AL108022 Drosophill
C 3	37.6	12.0	902	122	CNS006QP	AL065804 Drosophill
C 4	36.8	11.7	910	122	CNS006ON	AL065629 Drosophill
C 5	35.4	11.3	1149	122	CNS007PE	AL067457 Drosophill
C 6	35.2	11.2	1101	123	CNS0180I	AL108732 Drosophill
C 7	35.2	11.2	1201	122	CNS010A1	AL098707 Drosophill
C 8	35	11.1	1101	122	CNS00HZD	AL074538 Drosophill
C 9	34.8	11.1	1101	123	CNS016SL	AL107151 Drosophill
C 10	34.6	11.0	937	122	CNS006XP	AL066056 Drosophill
C 11	34	10.8	1101	122	CNS010E1	AL098851 Drosophill
C 12	33.8	10.8	844	122	CNS0052P	AL056652 Drosophill
C 13	33.8	10.8	1201	123	CNS01681	AL106411 Drosophill
C 14	33.6	10.7	284	64	AW148427	AW148427 xfl13h04.x
C 15	33.4	10.6	434	93	AQ058204	AQ058204 CIT-HSP-2
C 16	33.4	10.6	807	122	CNS0128R	AL101253 Drosophill
C 17	33.4	10.6	914	122	CNS00CVA	AL059884 Drosophill
C 18	33.4	10.6	1101	122	CNS002YD	AL098287 Drosophill
C 19	33.4	10.6	1184	123	CNS017XA	AL108616 Drosophill
C 20	33.2	10.6	462	79	AW634365	AW634365 bl19g11.w
C 21	33.2	10.6	518	105	AQ654017	AQ654017 Sheared D
C 22	33.2	10.6	520	42	A1664738	A1664738 TENG0714
C 23	33	10.5	465	91	W30090	W30090 mc26f03.r1
C 24	33	10.5	517	45	A1893152	A1893152 mc78f01.y
C 25	33	10.5	613	20	AA036271	AA036271 ml72d12.r
C 26	33	10.5	820	38	AI322633	AI322633 mj07b01.y
C 27	33	10.5	842	38	AI322989	AI322989 ml72d12.y
C 28	33	10.5	1101	123	CNS017X5	AL108611 Drosophill
C 29	33	10.5	1203	123	CNS015XR	AL106077 Drosophill
C 30	33	10.5	1225	123	CNS0166K	AL106358 Drosophill
C 31	32.8	10.4	301	30	AA767466	AA767466 ob46f03.s
C 32	32.8	10.4	649	26	AA440840	AA440840 LD15761.5
C 33	32.8	10.4	795	93	AQ048425	AQ048425 CLM-7f10-
C 34	32.8	10.4	1101	122	CNS00FKX	AL070921 Drosophill
C 35	32.8	10.4	1201	123	CNS015XQ	AL106040 Drosophill
C 36	32.6	10.4	417	113	AQ715927	AQ715927 HS_5430_B
C 37	32.6	10.4	1101	122	CNS007O	AL062090 Drosophill
C 38	32.6	10.4	1101	122	CNS00397	AL063912 Drosophill
C 39	32.6	10.4	1101	123	CNS016C5	AL107063 Drosophill
C 40	32.4	10.3	223	33	AA954528	AA954528 om96h03.s
C 41	32.4	10.3	1101	122	CNS00056	AL050564 Drosophill
C 42	32.4	10.3	1101	122	CNS00BNG	AL057398 Drosophill
C 43	32.4	10.3	1201	123	CNS0168S	AL106438 Drosophill
C 44	32.4	10.3	401	117	AQ0905303	AQ0905303 GSSTC10D1
C 45	32.2	10.3	440	39	AI450871	AI450871 ms01b06.x

ALIGNMENTS

RESULT	1
CNS0119V/c	
LOCUS	807 bp DNA
DEFINITION	GSS 26-JUL-1999
	CNS0119V
	Drosophila melanogaster genome survey sequence T7 end of BAC
	BACN03C05 of DrosBAC library from Drosophila melanogaster (fruit

ACCESSION	AL099997	fly), genomic survey sequence.			
VERSION	AL099997.1	GI:5611608			
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 807)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.				
FEATURES	Location/Qualifiers				
source	1..807				
	/organism="Drosophila melanogaster"				
	/plasmid="pBelOBAC11"				
	/db_xref="taxon:7227"				
	/clone_lib="DrosBAC"				
	/clone="BACN05C05"				
	/note="end : T7"				
BASE COUNT	157 a	38 c	17 g	287 t	308 others
ORIGIN					
Query Match	13.0%; Score 40.8; DB 122; Length 807;				
Best Local Similarity	6.0%; Pred.No. 0.1;				
Matches	12; Conservative 111; Mismatches 76; Indels 0; Gaps 0;				
QY	5	ccucgagagguuuaggaggaucgagcggcgagcgcgcgcgggaaguccaccuccgc	64		
Db	722	SSTKSSGGKGTAKKRGKKTGTGKGTGTGKTGTGKTSTGWSKBGTGKTBTBTBTDT	663		
QY	65	ucagcagggagcggcccgaucaucgucucuguguaucugauuguuugagccguc	124		
Db	662	TSTSTSTNTBTBTSTSTSTBKSSTTTSTKTSSTSTSTBTBTBTKTKTSTSTBT	603		
QY	125	agacggugauaauaagucguguguuuguguuuuuuuaccuuguguuuguc	184		
Db	602	DTTTSBTRGKWGKTGKGKGTAKBGTGKTGKTGTTTTTTTSTTKTTTBTSTTT	543		
QY	185	acuuugucagacgcgcgcug	203		
Db	542	TBTSTTSBSSBSBSSBS	524		
RESULT	2				
CNS017GS	1101 bp	DNA	GSS	26-JUL-1999	
LOCUS	Drosophila melanogaster genome survey sequence SP6 end of BAC				
DEFINITION	BACN17J04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108022				
VERSION	AL108022.1	GI:5628326			
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				

JOURNAL

JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr

COMMENT Determination of this BAC-end sequence was carried out as part of a

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebl.ac.uk>-. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

```

FEATURES
source      Location/Qualifiers
1. .1101

```

BASE COUNT ORIGIN	note-end : sp6	
268 a	126 c	195 t
	100 g	412 others

Query Match	12.5%;	Score 39.4;	DB 123;	length 1101;
Best Local Similarity	6.8%;	Pred. No. 0.26;		
Matches	15			

```

55 ccaaccuacgcagcagcagggagcccgaucaucagcagcucucuguguaucagaaugugug 114
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db YCCCCYBSSYBHYCCYCCYCCYCCYTTTKDRBWAHTTSTTBKTKTKTKTKTKTKTKTKTK 115
      856 YCCCCYBSSYBHYCCYCCYCCYCCYTTTKDRBWAHTTSTTBKTKTKTKTKTKTKTKTKTK

```

Db
916 KKKKKKSTSBKKBVGBBGGCGCTSBSSCCTKTCYBVCCTCTTCTCTKTTTAKKKTBT 975

Db 976 TKTAKKTTTGTGKKYBTTTYYKSBKKTKGYGKGGBYYKSSGYSKCSKCBSSCKCBY 103

Db 1036 GCBCTCYBKTTTTTKTKKCTKBTBTTKYYSKBYYKC 1074

RESULT 3

LOCUS	902 bp	DNA	GSS	03-JUN-1999
DEFINITION	CNS00606P			
DESCRIPTION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR14J09 of RPL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
FEATURES				
ORIGIN	1			
ORIGIN	2			
ORIGIN	3			
ORIGIN	4			
ORIGIN	5			
ORIGIN	6			
ORIGIN	7			
ORIGIN	8			
ORIGIN	9			
ORIGIN	10			
ORIGIN	11			
ORIGIN	12			
ORIGIN	13			
ORIGIN	14			
ORIGIN	15			
ORIGIN	16			
ORIGIN	17			
ORIGIN	18			
ORIGIN	19			
ORIGIN	20			
ORIGIN	21			
ORIGIN	22			
ORIGIN	23			
ORIGIN	24			
ORIGIN	25			
ORIGIN	26			
ORIGIN	27			
ORIGIN	28			
ORIGIN	29			
ORIGIN	30			
ORIGIN	31			
ORIGIN	32			
ORIGIN	33			
ORIGIN	34			
ORIGIN	35			
ORIGIN	36			
ORIGIN	37			
ORIGIN	38			
ORIGIN	39			
ORIGIN	40			
ORIGIN	41			
ORIGIN	42			
ORIGIN	43			
ORIGIN	44			
ORIGIN	45			
ORIGIN	46			
ORIGIN	47			
ORIGIN	48			
ORIGIN	49			
ORIGIN	50			
ORIGIN	51			
ORIGIN	52			
ORIGIN	53			
ORIGIN	54			
ORIGIN	55			
ORIGIN	56			
ORIGIN	57			
ORIGIN	58			
ORIGIN	59			
ORIGIN	60			
ORIGIN	61			
ORIGIN	62			
ORIGIN	63			
ORIGIN	64			
ORIGIN	65			
ORIGIN	66			
ORIGIN	67			
ORIGIN	68			
ORIGIN	69			
ORIGIN	70			
ORIGIN	71			
ORIGIN	72			
ORIGIN	73			
ORIGIN	74			
ORIGIN	75			
ORIGIN	76			
ORIGIN	77			
ORIGIN	78			
ORIGIN	79			
ORIGIN	80			
ORIGIN	81			
ORIGIN	82			
ORIGIN	83			
ORIGIN	84			
ORIGIN	85			
ORIGIN	86			
ORIGIN	87			
ORIGIN	88			
ORIGIN	89			
ORIGIN	90			
ORIGIN	91			
ORIGIN	92			
ORIGIN	93			
ORIGIN	94			
ORIGIN	95			
ORIGIN	96			
ORIGIN	97			
ORIGIN	98			
ORIGIN	99			
ORIGIN	100			

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 902)
Genoscope.
Submitted (02-JUN-1999) Genoscope - Centre National de Recherche

COMMENT

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

ECORI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP, P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	Location/Qualifiers
source	1 902

BASE COUNT	ORIGIN	/note- end : 1813-
298 a	50 c	61 g
		142 t
		351 others

Query Match	12.08;	Score 37.6;	DB 122;	Length 902;
Best Local Similarity	4.1%;	Pred. No. 0.85;		
Matches	8;	Conservative 126;	Mismatches 62;	Indels 0;
			Gaps	0

[illegible]

D**b** 751 KBITKTTTBTBCTBTBTBTTKYTYBTBTBTTBBBKYYTKBSBKBYTBSKBBYBT 810

Ddb 811 YVBTSSSTBBBNTYKTBTSBYCCTBYTBYBBBSTCBBSBSGSGSYSSSSSYBBSGBG 870

3271 TTTGGCTTATTCAGTCAACGCGA 3965

DD 8/1 JSTBKKBKBTBYSBKB 880

RESULT 4
CNS006ON

DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	210 bp	UNA	GSS	03-JUN-1999
ACCESSION	AL065629				

REFERENCE

1 (bases 1 to 910)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephypodroidea; Drosophilidae; *Drosophila*.

JOURNAL

COMMENT

Web : www.genoscope.cns.fr E-mail : segrete@genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of

melanogaster BAC library was prepared by Kazutoyo Oseogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the University of California, San Francisco.

P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

[illegible][illegible]

OY	80	cugaucugagcucuguguaucugauuuuguuggaaccgucuccagaacgugauaaau	139
Db	744	CCTTWTCTKCMCCGTTTTTTTCACMAAGCCCCSCCAATAGGGGGCGMCYCWGKDGMCAA	803
OY	140	aagucguguuuguguuuuuuuuuacccuuguguuugucacuuugucagacagcg	199
Db	804	AAGTTCCTTTTTTTTTTTTTTKTTHTCCTWAMTTTTTTTCGCGVMMGRASSGCSGKG	863
OY	200	ccugcgaaauugugucccacacgcgcggcuugcgaaauaacuuugagagucuuug	259
Db	864	SCYMCYSARAGGGGSGVCVCMAMGCMGGGCRGTGSTAAAVAAAAAACARTCYTTTTT	923
OY	260	ccuccagu	267
Db	924	TTKBMDT	931

RESULT	7
CNS010A1	
LOCUS	
DEFINITION	CNS010A1 1201 bp DNA GSS 26-JUL-1999
ACCESSION	Drosophila melanogaster genome survey sequence T7 end of BAC
VERSION	BACN03M23 of DrosBAC library from Drosophila melanogaster (fruit
KEYWORDS	fly)', genomic survey sequence.
SOURCE	AL098707
ORGANISM	AL098707.1 GI:5610318
	GSS.
	fruit fly.
	Drosophila melanogaster

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (23-JUL-1999) Genoscope.
1 (bases 1 to 1201)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope.

COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> . This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBeloBAC11.

```

FEATURES
source      Location/Qualifiers
1. 1201

```

BASE COUNT	ORIGIN	211 a	167 c	439 g	225 t	159 others
/note--end : T7"						

Query Match	11.28;	Score 35.2;	DB 122;	Length 1201;
Best Local Similarity	19.88;	Pred. No. 4.2;		
Matches	37;	Conservative	62;	Mismatches 80

```

QY      2   guaccucgcgaggguaauugggaggaucgcgagaguggcgggacgcucgcccgggaagcuaccaccuc 61
      1   | : | | | : : | : | | | | : : | : | | | : | |
Db      808  GTAGRRKGGRRAGGGGYICMGWACSGGGKGGGGGGGGGGRGKBTYGRGACGYACCAGAGR 86
      1   | : | | | | : : | : | | | | : : | : | | | : | |
QY      62   cgcucacgacaggggagacgccccugaucugagcucucuguguaucucgaaucuguauguaugacgcguc 121
      1   | : : | | | | : : | : | : | : | : | : | : : | :
Db      868  GSSSSCGCGGGGGRSGTKAALAAARRRRRGARGARAGRWGTGAGARRRRGAKTTTTTTTTTMT 927
      1   | : : | | | | : : | : | : | : | : | : | : : | :
QY      122  ccaagacguguaaauaauaagucguguguauguguauguauguaucuccuauguguauguc 181
      1   | : : : : : : : : : : : : : : : : : : : : : :
Db      928  CTTTTTTTTTTTTTTTWTMTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTSTYTKTTTTTTTK 987

```

QY	182	gucacuu	188
	:	1	:
Db	988	NTCTTTT	994

RESULT 8

LOCUS	DEFINITION	Accession
CNS000H2D	1101 bp DNA	GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC: BACR35A05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		AF074653.9

ACCESSION	AL074538	GI:49539144
VERSION	AL074538.1	
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	<i>Drosophila melanogaster</i>	

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

BASE COUNT	202	a	220	c	232	g	319	t	128	others
ORIGIN										
/organism="Drosophila melanogaster"										
/db_xref="taxon:7227"										
/clone_1b="Rpcl-98"										
/clone="BACR35A05"										
/note="end : T7"										

Query Match	11.1%;	Score 35;	DB 122;	Length 1101;
Best Local Similarity	23.8%;	Pred. No. 4.8;		
Matches 46;	Conservative	54.;	Mismatches	22

[illegible]

QY	7	ucgcagagguuuaggaggaucgcgaguggcggagacgcugcccggaagcuaccaccuccgcuc	66
		: : : : : :	
Db	1021	TTGKGKGKGGGGGGGGSGGGGCTKSCTGGGBGGGGGGCGGTCCGCTTSTYTTCC	1080
QY	67	agcaggggacgcgccugaucugagcugcuguguaucugaauguuuguaaccgucuccaag	126
		: : : : : : : : :	
Db	1081	CGTGGGGYYTCTTSCYTCCBCCTGGGKKTTTYTYTGTGKTBTCTNTBTBTKGK	1140
QY	127	acguguaaauaauaagucuguguuuguguaucuuacuuuguuugucgucac	186
		: : : : : : : : : : : : : : : : : : : : :	
Db	1141	KKTTTAKKGTGTGKSksNGGKGTGTTAKKKKKTAKKKKTGKKGTGTTAKKTBTC	1200
QY	187	u 187	
		:	
Db	1201	K 1201	

RESULT 14
AW148427/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW148427 284 bp mRNA EST 03-NOV-1999
xf13h04.x1 NCI_CGAP_Kid8 Homo sapiens cDNA clone IMAGE:2617975 3'
similar to contains element MER22 MER22 repetitive element ; , mRNA
sequence.
AW148427
AW148427.1 GI:6196323
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 284)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Mar 12, 1999 this sequence version replaced gi:4057126.

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from GIBCO
High quality sequence stop: 283.

FEATURES

Source

```

source
1. .284
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2617975"
/clone_lib="NCI_CGAP_Kid8"
/tissue_type="renal cell tumor"
/lab_host="DH108"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.2 kb. Life Technologies catalog #:
11524-014"

```

Query Match	10.7%	Score 33.6;	DB 64;	Length 284;
Best Local Similarity	31.2%;	Pred. No. 11;		
Matches 55; Conservative	32;	Mismatches 89;	Indels 0;	Gaps 0;

```

QY      26 ucgagagugcagcgacgcugcagcgaaagcuccaccuccgcucaagcagggagcgcgccuaguc 85
      : ||| ||| | ||| | ||| | ||| | :
Db     178 TTGGGGGGGGGGGGGGCCCGGGGGGGCCCCCCCCCCCCCGGGGGGGGGTGTGA 119

```

QY 86 ugaagcucuguguaucugauuguuuggaccgucuccaaagacgugauaanaaagucg 145
| | | | : : | | | | : : |
Db 118 ACCCCCCCGGGGTTTTTTTGGGGGGCCCCCGGGCCCGGGGGGGTTTTTTTGGGGG 59
| | | | : : | | | | : : |
QY 146 uguaauguguguuuguaucuaaguguuugucgucacuaugucagcgccc 201
| | : : : : : : : : : : : : : : | | | |
Db 58 GGGTTTTCTTTTTTTTTTTAGTTTTTTTTTTTTCCTCCCCCCCCC 3

RESULT	15			
LOCUS	AQ058204			
DEFINITION	AQ058204 434 bp DNA	GSS	30-JUL-1998	
ACCESSION	CIT-HSP-2345012	TR CIT-HSP	Homo sapiens	genomic clone 2345012
	genomic survey sequence.			
	AQ058204			

```

VERSION      AQ058204.1  GI:3354730
KEYWORDS     GSS.
SOURCE       human.
ORGANISM     Homo sapiens

```

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens	1 (bases 1 to 434)	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.	Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)	Unpublished (1998)	Other_GSSS: CIT-HSP-2345012.TF

Contact: Mark Palmer
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .434

Source

```

source
1. .434
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2345012"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBel0BAC11; site_1: HindIII; Site_2:
HindIII"
113 a      89 c      105 g      127 t
BASE COUNT
ORIGIN

```

Query Match	10.6%	Score 33.4	DB 93	Length 434
Best Local Similarity	28.0%	Pred. No. 13		
Matches	21	Mismatches	26	Indels 0
				Gaps 0

QY 101 cugaugungungagccgucucccaagacgugauanauaagucguguuugunguug 160
| : | : : | : | - : : | | | | : : | | | | : : | : | : |
DB 197 CTGAGTTTGCAGTTTAGTATCCAAAACATTCTGCACACAGCTTTTGTTGTGTTG 256

QY 161 uuuguuaccuuugu 175
: : : | : : :
DB 257 TTTGGTTGTTATAAT 271

Search completed: August 7, 2000, 11:56:12
Job time: 9746 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 10:55:16 ; Search time 4994.97 Seconds
(without alignments)
263.781 Million cell updates/sec

Title: US-09-214-124-2_COPY_452_578
Perfect score: 127
Sequence: 1 ugucgacagcgcccgcgaa.....ccucucgagcgagggauggg 127

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 segs, 5187315402 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_om: *
4: gb_om: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl1: *
8: gb_pl2: *
9: gb_pr1: *
10: gb_pr2: *
11: gb_pr3: *
12: gb_ro: *
13: gb_sts: *
14: gb_sy: *
15: gb_un: *
16: em_fun: *
17: em_hum1: *
18: em_hum2: *
19: em_in: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_sy: *
29: em_un: *
30: em_vl: *
31: gb_htg1: *
32: gb_htg2: *
33: gb_in1: *
34: gb_in2: *
35: em_ba1: *
36: em_ba2: *
37: em_hum3: *
38: em_hum4: *
39: gb_pr4: *
40: gb_htg3: *
41: gb_htg4: *
42: gb_htg5: *
43: gb_htg6: *

44: gb_htg7: *
45: em_htg1: *
46: em_htg2: *
47: em_htg3: *
48: em_hum5: *
49: gb_pl3: *
50: gb_pr5: *
51: gb_htg8: *
52: gb_htg9: *
53: gb_htg10: *
54: gb_htg11: *
55: gb_htg12: *
56: gb_htg13: *
57: gb_htg14: *
58: gb_in3: *
59: gb_htg15: *
60: gb_htg16: *
61: gb_htg17: *
62: em_htg4: *
63: em_htg5: *
64: em_htg6: *
65: em_htg7: *
66: em_hum6: *
67: gb_htg18: *
68: gb_htg19: *
69: gb_htg20: *
70: gb_htg21: *
71: gb_htg22: *
72: gb_htg23: *
73: gb_htg24: *
74: gb_htg25: *
75: gb_htg26: *
76: gb_htg27: *
77: gb_htg28: *
78: gb_htg29: *
79: gb_htg30: *
80: gb_htg31: *
81: gb_vl1: *
82: gb_vl2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	127	100.0	578	5	A83549	A83549 Sequence 2	
	2	127	100.0	940	5	A83548	A83548 Sequence 1	
	3	125.4	98.7	4292	82	AF006065	AF006065 Fowlpox v	
	4	120.6	95.0	1530	81	RESNVX	V01200 Spleen necr	
C	5	32	25.2	172078	40	AL160276	AL160276 Homo sapi	
C	6	32	25.2	187155	32	AL160271	AL160271 Homo sapi	
C	7	31.2	24.6	15089	4	GGRYR3	X95267 G.gallus mr	
C	8	30.2	23.8	3176	34	AF096275	AF096275 Drosophill	
C	9	30.2	23.8	27067	43	AC020041	AC020041 Drosophill	
C	10	30.2	23.8	165821	44	AC022005	AC022005 Homo sapi	
C	11	30.2	23.8	226318	34	AE003746	AE003746 Drosophill	
C	12	30	23.6	33556	43	AC020474	AC020474 Drosophill	
C	13	30	23.6	210788	11	CNS01DV7	AL133467 Human chr	
C	14	30	23.6	302915	34	AE003489	AE003489 Drosophill	
C	15	29.8	23.5	109016	8	ATT10K17	AL132977 Arabidops	
C	16	29.8	23.5	126569	73	AC008420	AC008420 Homo sapi	
C	17	29.8	23.5	214311	55	AC008834	AC008834 Homo sapi	
C	18	29.6	23.3	63377	42	AC018426	AC018426 Homo sapi	
C	19	29.6	23.3	103150	69	AC008720	AC008720 Homo sapi	
C	20	29.4	23.1	37243	33	LMFL685	AL138973 Leishmani	
C	21	29.2	23.0	538	3	BOVSEMI	M36982 Bovine semi	
C	22	29.2	23.0	2156	3	BTSPDNA	X62310 B.taurus se	
C	23	29	22.8	131967	55	AC016288	AC016288 Homo sapi	
C	24	29	22.8	186418	39	AC005871	AC005871 citb_109_	


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="q34.1-34.3"
/clone="RP11-323H21"
/clone_lib="RPCI-11.2"
37697 a 38143 c 37226 g 38069 t 36020 others
BASE COUNT
ORIGIN
```

OY	29	ccccacacgcgcgcgagcgaauaacuuugagagucuuuugccuccagugucucgcg	88
		: : : : :	
Dd	38229	CTCACCCTGGCATCATATGCTGTACTAAATTGGAGAGTGAATTGCCCTAAATGTGTCCTCC	38170
OY	89	uuuguacucgucucucucucucucucucgcg	116
		: : : : : : :	
Dd	38169	TTCCTGCTTGCCCTGCCTGCTCTCTG	38142

ACCESSION	X95267	GI:1212911
VERSION	X95267.1	
KEYWORDS	ryanodine receptor 3; RYR3 gene	
SOURCE	chicken.	
ORGANISM	Gallus gallus	

JOURNAL	Biochem. J.	201
MEDLINE	96207583	
REFERENCE	2 (bases 1 to 15089)	
AUTHORS	Sorrentino, V.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-JAN-1996)	
	V. Sorrentino, DIBIT, San Raffaele	
	Scientific Inst., via Olgettina 58, I-20132, Milano, ITALY	
	Location/Qualifiers	

```

/codon_start=1
/product="ryanodine receptor type 3"
/protein_id="CAA64563.1"
/db_xref="GI:1212912"
/db_xref="SPTREMBL:Q90985"
/translation="MAEGEGEGEDEIOFLRTDDEVVLQCVSSITHKEQRKFLAEGILG
NRLCFLEPTSEAKVPPDLCICNFVLEQSLSVRALQEMLTNTGDNASGAPAGGHRTL
LYGHAILLRHSFSEMYLTCLTSRSQCDKLAFDVGRLRENAAGEACWMTIHPAKORSE
GEKVRIGDILLIVSVSSERYLHLSMGSNIQDVASFMOTLWVNHPCSGSNTIEGYLL
GGHVRLFHGHDEYLLTPSTDONDQOOKVLYETGGAGVRAKSLMREYRLEKQDSTLK
IRMGQPFRLRHITGMVLYALNDEDEGLVMDREKSDTSSAFCFRASKELKEQHMDD
RDIDGMGVPEIKYGDVCFIQHVASALWLYKAPDAKSARLGLKRRVYLHQEHMDD
GLTLQRCQHEEQSAARIIRNTLSLFSQISGNRTLSPIALPVEEMAQTLQDLIKFYQ
PGEDLHEHDEKQKRLSLKRNQLFRKDEGMTALVNCIDRLNDYNSAAHFAEISRVEN
GTAWKEILNLLYELALIRGNRNNTQFSSNLDWLISKLRDLRESSGILEVYLHCILLI
ESPEALNIIAEHHIKSIIISLDKHGRNYKVLVDYCLSCVONGVAARANQLICDNLPL
RRDLLQTRLINDVTSIRPNISLIEGVAEGSAQYKRWYFELIIDQVDFLTAEPHLRVG
WASTSGVAPYPGGEGWNGNVGDDLVSFGFDGLHMSGRVPRAYASVNQHLSSDDVA
VSCCIDLGVPSISFRINGPVGQMFENFCTEGFFPEVYVSLAGVKAFLFLGRNGEERK

```

Location/Qualifiers
1. .187155

ELPPTGYAPCYEALLPKKMKLEPVKEYKRDSGVDLLGTOFLSQASFIPCPIDTS
QIALPEHLEKIRDKLAENIHELGMANKIELGWYGIKIRDDNKRHHPCLVFEFSKLPETE
KYNLQWSTETLTKTLALGCHIVHNPAAEEDLKXVKLPKNYIMSNKYKAPDLSEV
KLPSOEFLVNLKLAENAHNVAKDRIKQGWYGIQODLKNRNPRLVPYDLDBERTK
SNRDLREAVRTFAGYGVNVEPDQEIADQVEKVSIDKIRFYREVSQYAVKSGWYF
EEAVTGDMDRVGNAPRCRPIDELGADDAFVEEGSGKORWOGSGFGRSQWQGDV
VGCMLNDKSIIFTLNGELLITSGSELAFADEGIESGFVPIKSHIGLQIGRMNLT
DASTFYXTMCGLOEGFEPPAVNMNRDVAMFESKRLPTFVNPKNHPIEIMRIDGTI
ESPRLKVTHTLGTQNSNSDMIYCRISMPLEFRSSNFEGVGNASSDALOKRKHQ
EPPASSTCTCFSLRIFAGQDSSVWGWTPDYHESFNPDINKCTVTTLGDERGR
VHESVRSNCYMWGDIANSORSGRNVLDIEGCFVDLATGMLSTYANGKELGTCTY
OPEPNTKLFPAFVQPTSTNLIOELGKNTMPLSAIIFKSEERNPVQCPRLDVQ
TITPVLSRMPNSFLKVEVERSERHGVVOCLEPLOMALHIEENRCVDILELCEQ
EDLMKFHYHTLKLKLYSSVCLGNTRAVAYALCSHVDISQLEFYTDNOYLLPGLRSQFYL
LISIHLDHAKQAKLMNNNEFIIPYEEETRTIKLYPDETKKHGLPVGSLCKPSENF
STPCFIYSEHOTSSPDIPDLTKSAISMLTEAVOCSSHIRDPVGHIAFOFVY
LKLIALTLVGVFDDDDKQVLLIDPNVFGDKHEAEETEKEEVTQVEEKAVEAGE
KAVETKAPTGLIOTRLPESVKLOMCHLNTFCDELOHVEAIVSFADHYSKLOY
NOKYRNELMOALNMSAALTAKTKEFRSPQEQINMLNFQLEDCCPEEIRDEL
DFHDDLHICGIPLEEEEEEEDSLTGKRLSLIKIKPKPKPEKIEPREEDKSPTT
LKEHLSQTVRWSQEDIOPELVRIMYTLRQXDSIGELLQALRKAYITSAASVKD
TINLAALGOIRSLSVRMKEEELLMINGLGDIMNNKVFYOHPLMRVGMHETVMD
VMVNVLGDKSQIAFPKVAASCCRELFYFCRISRONKAMFEHLSTYLLNSSVGLAF
SMRSTPLDVAASVMDNDELALAEEDLDKVTYLAGCGLOSCPVLLAKGYDIPW
NPIEGERYLSFLRAFAVKNREGEDEEEEEEIYHGNAMSFYSALIDLGRCAPEM
EAIRISENPTRLPSOGYKRESEVEENASVVVKLLIRPECFGPRLRGSGNGLAAMQ
LIQSGKEAIRIRISIRSLVPTEDLVGISIPKLSTVNKDTGTVNEPDMASANCPDHA
APVLFELDRVYGIKDOSEFLHLEVGFLPDRASASLDVSLSTEALALNRYTCSA
VEPILKRCAPLPSGTEHSHASLVDSMPLKLTNNHYEQCWKYCLPSGMSYGAIAEDELH
PSMQOLRLRYLFDVPLNEXKMPKLTNNHYEQCWKYCLPSGMSYGAIAEDELH
LTEKFWGIFDSLSHKKYDPELEFPMALPCLSAIAGALPDYLDTRIRSTLEKQTSVP
EGNEDPKPVNTANVLPEKLEYIVSKYASHSHDKWAFDITNGKYGVSIDLETHTP
LIRPEKLTTEKEEIRYRWPRESLQWVEMAYENHNIMAKKMELESKGGSHPLLY
SSQNGYSPAPDLITNVLSRELQWVEMAYENHNIMAKKMELESKGGSHPLLY
PYDTLAKERSDRDKAQLKFLQVNGIISRGNDMDLASSMEKRAFKELKTL
KYVDSAQEFIAHLEAIVTSGTEKSPHDOIKFAKVLPLVDQYFTNHCYFLSSPT
KTLSSGAYNKEKEMVASFCKLAALVRHRISIFGSDSTMVSCHLIAOSLDRTY
MKSSELYKAGLRAFENAAEDLTSENKLGKFTHSRTQIKGVSONINTYVALP
VLTSEFHSIQYHFGVDLLGDVQVSCYRIICSLYSLGKNIVVERORPALCECLAS
LAAAIPIVAFLEPSLNHNPLSVNTKSAREAILGMDPTVEEMCEPIPOLDLIKEIN
NLASGARVTEMPHIVIEVILPMLCNLYSWMERGESYVESAGPCCTMTSHELSTIL
GNILKINNNGIDEASMMKRIAYVAPIISKARPDLKTHIPTLEKKAIAKIVM
EEBOLRADSKSDTOEALLIDEFAVLCRDLYAFYPMILIRYVDNNRANMLKPDADSD
ELFRVAAVEFILLWCKSHNFKREEQNFVIONEINNAFLTGDTSKMSKAMQVKSQGD
QERKSKRGDLYSIQTSILVAILKMLPIGLNCTPDGQELISLAKTRYSHRDTDEE
VEOPLRSKAVWHKLISKORRAVAVACFRMAPLYNLRHSINFLHGYQNYWIEET
YSEELKLVODLATSPKKEEEDTEKEQDPDLHQIILYFSRNALTEKLEDDPLYI
AYAAAMAKSCQEEEEEEDKTEFEKEMEKORTLYQARLHBRGAEMVLOMISAS
KGTGPMVETILKGIAILNGNTIQQKMDYLKLEKKGADGFEQSISGLMOSCVLDL
NAFEROSKAEGLMATEETLIVRERGEKVLQHDETRDLFRLOLCEGHNDFOQY
LRTQMGNTTNTNIIISTVDYLRLOESTISDFWYISGKVDIDESGORNESKALAVTKQ
IFNSLLEYIOGPCIGNOOSLAHSRLMDAVYVGLHVEANMQMSQDSAOIELLKELLD
LTKDMVVMLSLLEGNNVNGTIGKQMDTLVESPSNVELILKFDMLKLDLNTSDA
EKEHDPDGKTIISKDDQKSMEOQYIOSEIEFLSCAEDENDMNYIDFEVERHE
PAKDIGFNNAVLLTNSEHMPNDSRLQSLLEPAEESVLYNFEPIGRIEIMGAKKIE
RYFEISESRRMQWEKPOVKESKROFIEGIVNEGGEEMKELFVNCEDTIFETOLAS
QISESDSRELRKIEEPCYVVIDIDDEEEKSLESPSAFALACAAYKKNVANEKMYT
VKNLRKOYRKVRKMTVKEMVPCSLFEWILFVGFQLEFTIIVGIFQILMSTVEGGL
VEGAKNIKVTKIIGDMPDPTOGIHDVTEAKTEGAEHIGIRDELVOFVKEGETDI
ISDIFGIPSKKEGSGKHGDAIGDIAEILGSDIOSLENNARKKGLQIPEAKAE
AERKVEAEKADMDGEKODKAKEEQSOPEEGKMKRRRKHQKIEKPEAVMANFEKA
LEIYQTKILHYLARNFYLNELFALFVAFAINELLEYKVTBEPDEVEEDSNLMSFE
EEEBEGMVFVLEESTGYMAPALRALALFIPISFVCIYGYCLKVPAYEKRREKEV
ARKLEFDGLYTEOPSEDDIKGOWDRLVINTPSFPNNWDKFVKRVINKYGDLYGAE
RIAEELGLDKNARHFCPEEPEASLSVSSIDTKYHIMKLGVTDNSEFLYLAW
YTTMSILGHYNNFRAHLDDIAMGFKTLRTIISVTHNGKQLVLYGILLAVVLYLT
VVAFNEFRKFYKNSDEDEPDMKCDMMCYLFHMYGVGRAGGIGIDEIEDPADPYE
MYRIAFDITFEFFVYVILLAIIOGLIDAFGLRDQEQVREDMETKCFICIGINDYF
DTIPHGETHTLQEHNLANYLFLMYLINKDETEHTGQESFVWKMQERCWDFEPAG
CFRKQYEDQLG"

BASE COUNT 4713 a 2814 c 3374 g 4186 t 2 others
ORIGIN

Query Match 24.6%; Score 31.2; DB 4; Length 15089;
Best Local Similarity 32.4%; Pred. No. 12;
Matches 22; Conservative 23; Mismatches 23; Indels 0; Gaps 0;
QY 54 uacuuugagagucuuuugccuacugucuuucguuuuugacucugucucucucucucuc 113
Db 10921 TGCTTTCCATTCTTCTTCTCCTCAATGCTTCTTCTTGTCTTCTCCTCCTCTTCTT 10862
QY 114 ccgagccgg 121
Db 10861 TCTTCTG 10854

RESULT 8
AF096275/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3176)
McColl, G. and McKechnie, S.W.
The heat shock gene hsp68 of D. melanogaster
Unpublished
2 (bases 1 to 3176)
McColl, G. and McKechnie, S.W.
Direct Submission
Submitted (02-OCT-1998) Department of Biological Sciences, Monash
University, Wellington Rd, Clayton, Vic 3168, Australia
Location/Qualifiers
1. 3176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3"
/map="95D5-D9"
/1157..>3064
/gene="hsp68"
/product="heat shock protein 68"
/1157..>3064
/gene="hsp68"
/allele="hsp68d"
1157..3064
/gene="hsp68"
/codon_start=1
/product="heat shock protein 68"
/protein_id="AADI6140.1"
/db_xref="GI:4322942"
/translation="MPAIGIDLTGTYSCVGVFQYKVEIANDQGNRTTPSYVAFETDS
ERLIGDAKNQVAMNPKNSVFDAKRLIGRFRDSDSKIQEDIKHWPFKVINDNKPKRISV
EFGKANGCFSPPEIISSMVLTQKMETAEAYLGTGVKDAVITVPAYFNDQROAKDAGA
IAGINVLRIINEPTAAALAYGLDKNLKGERNVLFIDLGCTFDSVILTIDEGSLEEV
STAGDTHLGGEDFDNRLVNHAEFEKRYKKDLRSNPRALRLRTAAERAKRTLSST
EASLEIDALYEGHDFYSKYSARFEELGGLFRNTELEPVEKALKDAKMDKSIHDIIVL
VGGSTRIPKVNQLONFEGKTLNLSINPDEAVAYGAIIQAAIISGDKSSSEIKDVLV
DYAPLSIGIETAGVNTKLIERNRIIPCKOSKFTTYADNOPAVTIQVEGERALTKD
NNVIGTFDLTPPAPRGVPKIDVTFDLANGILNTAKOQGTGNAKNTITKNDKRL
SOADIDRLMSEAKEYAEEDERHRQRIARNOLEFYLEGVKAAENGSDRTISAADKSSI
VERCEAMKWLDSNTTAEKEEYKLELEDFQSPIMTKMHKGGGCGQAAPNFGQAQAG
GYKGPVTEVVD"

BASE COUNT 965 a 744 c 721 g 746 t
ORIGIN
Query Match 23.8%; Score 30.2; DB 34; Length 3176;
Best Local Similarity 47.1%; Pred. No. 27;

Matches	24;	Conservative	14;	Mismatches	13;	Indels	0;	Score	100
---------	-----	--------------	-----	------------	-----	--------	----	-------	-----

Db 2943 TCCAGTTCCTCAGTTGTACTCGTACCTCCCTCCCTTCGCGGGGAGGGATGC

RESULT	9	HTG	03-JAN-2000
AC020041/c			
LOCUS	27067 bp	DNA	
DEFINITION	AC020041	*** SEQUENCING IN PROGRESS ***	in ordered
	Drosophila melanogaster,		
	bioass		

ACCESSION	AC020041	GI:6664856
VERSION	AC020041.1	
KEYWORDS	HTG; HTGS_PHASE2.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
	metazoa: arthropoda:	Tracheata; Hexapoda; Insecta;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 27067)	Adams, M. and Venter, J.C.	Direct Submission	Celera Genomics, 45 West Gude Drive, Rockville, MD, USA	This sequence was identified as CDM:10211839 by the submitter. Information on this record e-mail to fly@celera.com.

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

```

FEATURES
  source
    LOCATION/GenBank
    1. .27067
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
BASE COUNT
7813 a 5994 c 5686 g 7574 t
ORIGIN

```

Query Match	23.88;	Score 30.2;	DB 43;	Length 27067;
Best Local Similarity	47.18;	Pred. No. 23;		
Matches 24;	Conservative 14;	Mismatches 13;	Indels 0;	Gaps 0;

QY 75 uccagugucuuccguuuuacucgucuccucuccuuccucgcggcggaug 125
:||||| :||| :|||:||||| :|||:| ||| | | :|
Db 17023 TCCAGTTCCTCAGTTGTACTCGTACTCCTCCTTCTCGGCGGTGGTGTTG 16973

	RESULT	ID	HTG	DATE
AC0222005/c	10			03-FEB-2000
LOCUS		165821 bp	DNA	
DEFINITION		Homo sapiens chromosome 3p clone RP11-378N17,	WORKING DRAFT	
		SEQUENCE. 19 unordered pieces.		

ACCESSION	AC022005	GI:6742900
VERSION	AC022005.1	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT	
SOURCE	human.	
ORGANISM	Homo sapiens	Chordata

ORGANISM	Homio sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE	1 (Bases 1 to 165821)	
AUTHORS	Wang, J., Hu, S., Dong, W., Zhang, X., Wang, X., Zhang, Y., Zhang, H., Wu, O., Wang, H., Yang, X., Cheng, C., Wei, Y.	

TITLE	REFERENCE	AUTHORS
Chromosome 3p genomic sequence unpublished	2 (bases 1 to 165821)	Zhang, J., Hu, S., Dong, W., Zhang, X., Wang, X., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Ni, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H.,

TITLE
 Lǐu, X., Lǐ, G., Lǐ, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L.,
 Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L.,
 Feng, X., Yu, J. and Yang, H.
 Direct Submission
 Submitted (24-JAN-2000) Human Genomic Center, Institute of
 Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 JOURNAL

100101, P.R.China 'working draft' sequence. It currently
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 19 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

[illegible]

```

FEATURES
  source
    location/quantity
      1. 165821
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="3p"
        /clone="RP11-378N17"
    54104 a 30480 c 30787 g 50446 t 4 others
ORIGIN

```

Query Match	23.8%;	Score 30.2;	DB 44;	Length 165821;
Best Local Similarity	35.0%;	Pred. No. 19;		
Matches 43; Conservative	22;	Mismatches	58;	Indels 0; Gaps 0

QY 1 ugucgacagcgcccuugcgaauugugucccacaccgcgcgcguugcgaauaacuuug 60
 : |^a| ||| ||| : | ::||| | ||||| ::| ||| ||| | : |
 Db 13903 TGACCACAAATGCCATTGGTTTGTGCAGACCACAATTTATCATCTGAGAAGAACATG 13844
 : |^b| ||| ||| : | ::||| | ||||| ::| ||| ||| | : |

QY	61	gagagucuuuugccuccagugucuuuccguuuuguaucucgucucucucuccucucg	120
Db	13843	GTGAACATTTTCATCGAGTGTTCCTAGATGAAACTGGTTATGCTTTTTCAG	13784

OY 121 gga 123
Db 13783 GGA 13781

RESULT 11

AE003746

LOCUS

DEFINITION

AE003746 226318 bp DNA INV 22-MAR-2000
Drosophila melanogaster genomic scaffold 142000013386035 section 71
of 105, complete sequence.

AE003746 AE002708

AE003746.1 GI:7301086

HTG.

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 226318)

REFERENCE

AUTHORS

ADAMS, M.D., CELNIKER, S.E., HOLT, R.A., EVANS, C.A., GOCAYNE, J.D.,
AMANATIDES, P.G., SCHERER, S.E., LI, P.W., HOSKINS, R.A., GALLE, R.F.,
GEORGE, R.A., LEWIS, S.E., RICHARDS, S., ASHBURNER, M., HENDERSON, S.N.,
SUTTON, G.G., WORTMAN, J.R., YANDELL, M.D., ZHANG, Q., CHEN, L.X.,
BRANDON, R.C., ROGERS, Y.H., BLAZEJ, R.G., CHAMPE, M., PFEIFFER, B.D.,
WAN, K.H., DOYLE, C., BAXTER, E.G., HELT, G., NELSON, C.R., GABOR
MIKLOS, G.L., ABRIL, J.F., AGBAYANI, A., AN, H.J.,
ANDREWS-PFANKOCH, C., BALDWIN, D., BALLEW, R.M., BASU, A.,
BAXENDALE, J., BAYRAKTAROGLU, L., BEASLEY, E.M., BEESON, K.Y.,
BENOS, P.V., BERMAN, B.P., BHANDARI, D., BOLSHAKOV, S., BORKOVA, D.,
BOTCHAN, M.R., BOUCK, J., BROKSTEIN, P., BROTTLER, P., BURTIS, K.C.,
BUSAM, D.A., BUTLER, H., CADDEU, E., CENTER, A., CHANDRA, I.,
CHERRY, J.M., CAWLEY, S., DAHLKE, C., DAVENPORT, L.B., DAVIES, P., DE
PABLOS, B., DELCHER, A., DENG, Z., MAYS, A.D., DEW, I., DIETZ, S.M.,
DODSON, K., DOUP, L.E., DOWNS, M., DUGAN-ROCHA, S., DUNKOV, B.C.,
DUNN, P., DURLIN, K.J., EVANGELISTA, C.C., FERRAZ, C., FERRIERA, S.,
FLEISCHMANN, W., FOSLER, C., GABRIELIAN, A.E., GARG, N.S.,
GELBART, W.M., GLASSER, K., GLODEK, A., GONG, F., GORRELL, J.H., GU, Z.,
GUAN, P., HARRIS, M., HARRIS, N.L., HARVEY, D., HELMAN, T.J.,
HERNANDEZ, J.R., HOUCK, J., HOSTIN, D., HOUSTON, K.A., HOWLAND, T.J.,
WEI, M.H., IBEGWAM, C., JALALI, M., KALUSH, F., KARPEN, G.H., KE, Z.,
KENNISON, J.A., KETCHUM, K.A., KIMMEL, B.E., KODIRA, C.D., KRAFT, C.,
KRAVITZ, S., KULP, D., LAI, Z., LASKO, P., LEI, Y., LEVITSKY, A.A.,
LI, J., LI, Z., LIANG, Y., LIN, X., LIU, X., MATTEI, B., MCINTOSH, T.C.,
MCLEOD, M.P., MCPHERSON, D., MERKULOV, G., MILSHINA, N.V., MOBARRY, C.,
MORRIS, J., MOSHREFI, A., MOUNT, S.M., MOY, M., MURPHY, B., MURPHY, L.,
MUZNY, D.M., NELSON, D.L., NELSON, D.R., NELSON, K.A., NIXON, K.,
NUSKERN, D.R., PACLEB, J.M., PALAZZOLO, M., PLITMAN, G.S., PAN, S.,
POLLARD, J., PURI, V., REESE, M.G., REINERT, K., REMINGTON, K.,
SAUNDERS, R.D., SCHEELER, F., SHEN, H., SHUE, B.C., SIDEN-KLAMOS, I.,
SIMPSON, M., SKUPSKI, M.P., SMITH, T., SPIER, E., SPRADLING, A.C.,
STAPLETON, M., STRONG, R., SUN, E., SVIRSKAS, R., TECTOR, C., TURNER, R.,
VENTER, E., WANG, A.H., WANG, X., WANG, Z.Y., WASSARMAN, D.A.,
WEINSTOCK, G.M., WEISSENBACK, J., WILLIAMS, S.M., WOODAGE, T.,
WORLEY, K.C., WU, D., YANG, S., YAO, Q.A., YE, J., YEH, R.F.,
ZAVERT, J.S., ZHAN, M., ZHANG, G., ZHAO, Q., ZHENG, L., ZHENG, X.H.,
ZHONG, F.N., ZHONG, W., ZHOU, X., ZHU, S., ZHU, X., SMITH, H.O.,
GIBBS, R.A., MYERS, E.W., RUBIN, G.M. and VENTER, J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)

TITLE
JOURNAL
MEDLINE
REFERENCE

AUTHORS
TITLE
JOURNAL

FEATURES
source

MRNA
Location/Qualifiers
1. 226318
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
join(<9315. .9453,9516. .9809,9878. .>11271)
/gene="CG5933"

gene

CDS

MRNA

gene

CDS

MRNA

gene

CDS

/product="CT18635"
/db_xref="FLYBASE:FBan0005933"
/db_xref="FLYBASE:FBgn0039139"
/gene="CG5933"
/db_xref="FLYBASE:FBan0005933"
/db_xref="FLYBASE:FBgn0039139"
/gene="CG5933"
/note="CG5933 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0005933"
/db_xref="FLYBASE:FBgn0039139"
/protein_id="AAF56221.1"
/db_xref="GI:7301087"
/translation="MADAMDIKSLKTRNTLREKLEKRRKEREIILSDIOEDLTNPKK
ELVEADLEVQKEVLQALSSCSIALPIVSTQVVEKIASSSLEWVFILGKLANOGAIVI
RNVTIGTEAGCEIISVQPKELKEILEDTNDTCQOKEEAKRLEVDVDOPOEKTIRL
ESTVARKESTSDAPDDIMMLTSMPTREKOSKOVGEILELITKPTAKERSVAEKER
KTESAEDEGEIIAEVINNCEAESESTGSDTCSSETDCKTLHFKKIIQAHTEBSL
GDCSFLNCFHMACTCKVYHVEVDLPHINTNPTDVKTLSLRSVDSCTLYPPQWI
OCDRLFLNMTVLGKFVAVVADPMDIMELPYGTMSDDEKALGVPALODDGLFLWV
TGRAMELGRCDCIKMGYERVELIWKTNQLORIIRTGRLGHLNKGHEHCLVGMGN
PTNLNRGLDGDVIVAEVRATSHKDEIYGLIERLSPTRIEELFGRPHNIQPNWITLG
NOLDGIRLVDPELITQFOKRYPDGNCMSPASANAASINGIQK"
complement(join(12728. .13936,14049. .>14921))
/gene="CG5410"
/product="CT17166"
/db_xref="FLYBASE:FBan0005410"
/db_xref="FLYBASE:FBgn0039140"
complement(<12728. .>14921)
/gene="CG5410"
/db_xref="FLYBASE:FBan0005410"
/db_xref="FLYBASE:FBgn0039140"
complement(join(12728. .13936,14049. .14666))
/gene="CG5410"
/note="CG5410 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0005410"
/db_xref="FLYBASE:FBgn0039140"
/protein_id="AAF56222.1"
/db_xref="GI:7301088"
/translation="MGQYTASQKRNVRILLVGDAVGKTSILISLVSESEYPEVPRA
EETIPANVTPCOVPTSIDVESAVEQSEDALAEINKAHVVCIVYAVDDDDTLDRITS
HWLPLVRAKNPSLDGEGDAEAEAGDTQREPIRKPIVLGNKIDILIEYSTMSVLAI
MEDYPEIESCVCESAKSLHNISEMEFYAQAVALHPTSPLYMNEQELTSACKSLVRI
FKICDIDGNDLNDYELNLFQRCFNTPLQPOLDEVKAVIQKNPVDGIYNDAVITLKG
FLFLHCLFIQGRNETTWAVLRFGYNDQLEMCQEYLRPPKIPGSSSTELSHRGQOF
LIAVFERVDRDGDGALSPEEHKMFSTCPAAPWSYSTDIRKSCPINETGTWVTLHGM
CRWTIMTLIDVVKMETEYLAIFGVNHNENDSQLAAIHVTRERRIDLAKROSSRSYVKH
VIGPKSGKTMCGRGLVEDMHLIKGFEKTNVNCINSYQVYQGEKHLILRDIDVRH
ALDPLQOEVCNDAVACLVDSNPRSFEYVARIYIKIYAESKIPVMIVGTCKDMERR
QDYLMQSEFCDKYKLLPPLHLSLTKNKELTYKLATMAAFP"
join(15341. .16477,16713. .16883,17029. .17214,17424. .17590,
17686. .17938,17997. .18742)
/gene="CG5977"
/product="CT18787"
/db_xref="FLYBASE:FBan0005977"
/db_xref="FLYBASE:FBgn0039141"
/gene="CG5977"
/db_xref="FLYBASE:FBan0005977"
/db_xref="FLYBASE:FBgn0039141"
/note="CG5977 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0005977"
/db_xref="FLYBASE:FBgn0039141"
/protein_id="AAF56223.1"
/db_xref="GI:7301089"
/translation="MVRTKNQSSSSASSSSTKSPIKSSSGAGSSGCGLGROSTHRS

TITLE Direct Submission
JOURNAL Submitted (19-JAN-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cremlieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr

COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
Location/Qualifiers
1. .109016
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
1. .5127
/note="overlap to BAC F15B8, please refer to Acc no. EMBL:AL049660 for analysis and annotation"
3113. .6400
/gene="T10K17.10"
complement(3113. .3385)
/number=1
complement(join(3113. .3385,4086. .4157,4306. .4374,5512. .5577,5669. .5782,5855. .6400))
/note="similarity to several hypothetical proteins - Arabidopsis thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CAB67608.1"
/db_xref="GI:6729523"
/translation="MDLTGGFGARSGVGPCREPIGLESLHGDDEFROLVTTLPENP GGSFTALBELPTQAVELHFTDSSSSQAAVTGIGEIPPLHSFGTILAFPSNVL MERAARFSVIATEQONGNISGETPTSSVPSNSANLDRVKTETPAETDSSQRLISDSAI ENQIPCPNQNRRNGKRDKEFGKSSSTKKNSSENEKLPYVHVRARGOATDSHSLA ERARREKINARMKLOELVPGCDKIQTALVDEIINHVSQLOQVEMLSMRLAAVNP RIDFNLDITILASENSLMDGSFNAPMOLAVPQQAIEQSFHHRQLQQPPTQWPFPD GLNOPVMGREDDOAHGNDNSNLMAVSENVAVSANLHPNQVKMEL"
complement(3386. .4085)
/number=1
complement(4086. .4157)
/number=2
complement(4158. .4305)
/number=2
complement(4306. .4374)
/number=3
complement(4375. .5511)
/number=3
complement(5512. .5577)
/number=4
complement(5578. .5668)
/number=4
complement(5669. .5782)
/number=5
complement(5783. .5854)
/number=5
complement(5855. .6400)
/number=6
9838. .109016
/note="overlap to BAC F9D24"
11415. .11928
/gene="T10K17.20"
/number=1
11415. .12888
/gene="T10K17.20"
join(11415. .11928,12048. .12130,12308. .12348,12502. .12600,12711. .12888)
/gene="T10K17.20"
/codon_start=1
/product="putative protein"
/protein_id="CAB67609.1"
/db_xref="GI:6729524"

/translation="MMLCYSPITTCRNAISIKRHLGSRLYGVVAHGSSKFCYSLLS GLSRHYTGFRVSVSNRPSWHDKGLFGSVLINRPTVAPKEKLEVSFLSPKANMCKSK IESNMKNLYCYSRFAYTGIVISLVYSSTSQSAVADSSRDKDANNVHHSSDGKFPN GKRYVTDYSIIAHGFCLSGSKLAPGEKMORELDELRTVRADVIQROTEWFEVGD EDTVVRQIRDPHVWGGEPEPELMASHVLOMPITVYMKDDKAGGLISIAEYGOEYKDDP IRLVHGFGHYDALLLHESKASIPKSKL"
11929. .12047
/gene="T10K17.20"
/number=1
12048. .12130
/gene="T10K17.20"
/number=2
12131. .12307
/gene="T10K17.20"
/number=2
12308. .12348
/gene="T10K17.20"
/number=3
12349. .12501
/gene="T10K17.20"
/number=3
12502. .12600
/gene="T10K17.20"
/number=4
12601. .12710
/gene="T10K17.20"
/number=4
12711. .12888
/gene="T10K17.20"
/number=5
13360. .13903
/gene="T10K17.30"
complement(join(13360. .13528,13803. .13903))
/note="ESTs matching in this region do not correspond to any gene models eg.GB:AA651588; similarity to 60S RIBOSOMAL PROTEIN L21 - Arabidopsis thaliana, SWISSPROT:RL21_ARATH"
/codon_start=1
/product="putative protein"
/protein_id="CAB67610.1"
/db_xref="GI:6729525"
/translation="MHVNIPLMSPTTVPDGIHTGSKQAYPSFTGLKWVYQIGKII RKRIHVLVEHYOOSRCAVEFKLRKKNDLKAASKARGETISTKR"
complement(13360. .13528)
/number=1
complement(13529. .13802)
/number=1
complement(13803. .13903)
/number=2
14857. .17399
/gene="T10K17.40"
14857. .16195
/gene="T10K17.40"
/number=1
join(14857. .16195,16750. .17399)
/gene="T10K17.40"
/note="strong similarity to several receptor-like protein kinases"
/codon_start=1
/product="receptor-like protein kinase"
/protein_id="CAB67611.1"
/db_xref="GI:6729526"
/translation="MQLRLTLTVSSYFEFFINESSSLNPDGLSLALKSAILRDPY RVMTWSWESDPTPCWHPGIICTHGRVTSVLVSGRRLSGYIPSKGLIDSLIKLDIARN NESKPVPTRLFNAYNLRYIDLSHNSISGPIPAQIOSLKNLTHIDSSNLNGSLPOSTL TQGLSVGTNLNSYNSFSEGEIPSPYGRFPYVSLDLGHNNLTGKIPQIGSLNQGPTA FAGNSELCGFPLOKCKDEGTNPKLVAPREGSOILPKKPNPSPFDKDGKRNKPITGS VTVSLISGVSIYGAIVSISVWLIRKLSSTVSTPEKNNTAAPLDDAADDEEKEGKFPV MDEGFLELEDLIRASAYVVGKSRSGIYRVVAGMGSGTVAATFTSSTVAVRRLSDG DATWRKRDPFENEVEAISRVQHPNIVRLRAYYYAEDERLITDIYIRNGSLYSALHGSPS NTLPSLSWPERLLIAQGTARGLMYIHEYSPPRYVHGNLKSKILLDDDELLPRISGFGI TRLVSGYSKLGISLATSROSIDQTYLTSATTVTRITAPYAVYLAPEARASSGCKLSOK CDVYSFGVYLMELLTGRLLPNASSKNNGEELVRVVRNWKEEKPLSEILDPEILNKGHA


```

DKOVIAAIHVALNCTEMDPEVPRPMRSVSESLGRIKSD"
16196. .16749
/ gene="T10K17.40"
/ number=1
16750. .17399
/ gene="T10K17.40"
/ number=2
17902. .18366
/ gene="T10K17.50"
17902. .18366
/ gene="T10K17.50"
/ number=1
17902. .18366
/ gene="T10K17.50"
/ codon_start=1
/ product="hypothetical protein"
/ protein_id="CAB67612.1"
/ db_xref="GI:6729527"
/ translation="MASQRKLIMVILSLMKVALSQQYGVMGKSIFKFEFFPMISY
YITNDIGGGLVHSGCITSRNGYRRIRDFPGSMKIFAERKTYGRTYHCEFFRGD
ETQIHRSFKDIRNIDKYYQCRHCFWSIRNPGCALNSHTGKYDICYAMDK"
18373. .18408
/ note="36 bp TA tandem repeat"
19456. .19870
/ gene="T10K17.60"
19456. .19543
/ gene="T10K17.60"
/ number=1
join(19456. .19543,19686. .19870)
/ gene="T10K17.60"
/ codon_start=1
/ product="hypothetical protein"
/ protein_id="CAB67613.1"
/ db_xref="GI:6729528"
/ translation="MASERKLIMVILSLMKVALSQQYGVMGDYSFNSGKGFGEEL
AMFVNLDGMKDKYVILLFIKITEIIFTNMSADTVFGPSEGMDLVP"
19544. .19685

intron

Query Match      23.5%; Score 29.8; DB 8; Length 109016;
Best Local Similarity 31.5%; Pred. No. 27;
Matches 23; Conservative 23; Mismatches 27; Indels 0; Gaps 0;

QY      43  cuugcgauaauacuugagagagucuuuugccuccagugucuccguuugacuugucu 102
||:| || ||:|:|:| ||:|:|:| | : | : |:|:|:| : |:|
Db      41692  CTTTCTAAATACTTAGTGATCTTTCTTTTAAGCAATGAGATTGTCTTCTTCT 41751

QY      103  ccuuccccuucc 115
||:| || | : |||
Db      41752  TCTCACCATAATCC 41764

```

Search completed: August 7, 2000, 10:56:58
Job time: 6192 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2000, 09:21:37 ; Search time 205.61 seconds
(without alignments)
154.537 Million cell updates/sec

Title: US-09-214-124-2_COPY_452_578
Perfect score: 127
Sequence: 1 ugucgacagcgcccgccgcaaa.....ccucucgagcgagggauggg 127

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 segs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	127	100.0	578	1	V63779	Sequence derived f
2	127	100.0	940	1	V63778	Sequence derived f
3	125.4	98.7	4643	1	X26113	Fowlpox virus (FPV)
4	116.6	91.8	5519	1	Q76040	Retrovirus vector
5	116.6	91.8	5525	1	Q76039	Retrovirus vector
6	116.6	91.8	5528	1	Q76038	Retrovirus vector
7	113.6	89.4	3878	1	Q76041	Retrovirus vector
8	28.8	22.7	3946	1	Q80523	Human mcl-1 gene.
9	27.4	21.6	333	1	T11309	Allergen Alt a 12
10	27	21.3	2487	1	Q45332	Human epidermal su
11	27	21.3	2488	1	T95887	DNA for epidermal
12	27	21.3	49998	1	X23518	Human kidney amino
13	26.8	21.1	1760	1	V38240	Human liver activi
14	26.2	20.6	1335	1	Q81700	C2P2(75-520). New
15	26.2	20.6	2216	1	Q70072	Canine zona pelluc
16	26.2	20.6	2381	1	Q65608	Canine zona pelluc
17	26.2	20.6	2381	1	V64790	Canine ZPA CDNA. I
18	26	20.5	2389	1	T62793	Mouse knos CDNA. K
19	25.8	20.3	3919	1	X13618	Enterococcus faeca
20	25.6	20.2	312	1	T22914	Human gene signatu
21	25.6	20.2	1206	1	T10184	Trypsin-like prote
22	25.4	20.0	2040	1	T62137	Leishmania brazili
23	25.4	20.0	2040	1	V47557	Leishmania antigen
24	25.4	20.0	3899	1	V63192	CDNA:from clone dt
25	25.4	20.0	15464	1	Q11415	Ryanodine receptor
26	25.2	19.8	2419	1	T16883	Human growth diffe
27	25.2	19.8	7555	1	Q05831	Cardiac sodium cha
28	25.2	19.8	7555	1	Q81328	Cardiac sodium cha
29	25	19.7	1674	1	X22532	Human MMP-20 DNA.
30	25	19.7	2017	1	T60713	Saccharomyces cere
31	25	19.7	7393	1	T74883	Porcine retrovirus
32	25	19.7	8196	1	V09699	Porcine retrovirus
33	25	19.7	8209	1	V09700	Porcine retrovirus

C	34	24.8	19.5	156	1	V07081	DNA encoding HIV-1
C	35	24.8	19.5	2301	1	V20445	Human c-erbB onco
C	36	24.8	19.5	2304	1	X05857	Rat pheromone rece
C	37	24.8	19.5	3584	1	X05818	Rat pheromone rece
C	38	24.8	19.5	7159	1	V74613	Staphylococcus aur
C	39	24.6	19.4	2089	1	T05513	Murine neurogenic
C	40	24.6	19.4	2089	1	T74887	Mouse neurogenic d
C	41	24.6	19.4	2089	1	V42928	CDNA encoding mur
C	42	24.6	19.4	2172	1	V02996	Mammalian Ena (Men
C	43	24.6	19.4	2898	1	V02996	Mouse neural Menat
C	44	24.6	19.4	3075	1	V06252	Murine puromycin-s
C	45	24.6	19.4	3710	1	V74716	Staphylococcus aur

ALIGNMENTS

RESULT 1	
ID V63779	standard; DNA; 578 BP.
AC V63779;	
DT 17-MAR-1999	(first entry)
DE Sequence derived from the 5' end of REV-A.	
KW Reticuloendotheliosis virus type A; REV-A; 5' end genomic RNA;	
KW type C retrovirus; internal ribosome entry site; IRES; encapsidation;	
KW retroviral vector; gene therapy; ss.	
OS Reticuloendotheliosis virus.	
PN FR2762615-A1.	
PD 30-OCT-1998.	
PE 28-APR-1997; 005203.	
PR 28-APR-1997; FR-005203.	
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.	
PI Darlix JL, Gabus Darlix C, Lopez LM;	
DR WPI; 99-037487/04.	
PT Expression vectors containing IRES and/or encapsidation enhancer -	
PT derived from type C retrovirus other than FMLV and MoMLV	
PS Claim 6; Page 32; 43pp; French.	
CC The present sequence is derived from the 5' end of the genomic RNA of	
CC Reticuloendotheliosis virus type A (REV-A). The specification also	
CC describes nucleotide sequence derived from all or part of the 5' end	
CC of the genomic RNA of a type C retrovirus other than Friend murine	
CC leukaemia virus (FMLV) and Moloney murine leukaemia virus (MoMLV). The	
CC 5' derived sequences are used as an internal ribosome entry site (IRES)	
CC in a vector for permitting or enhancing the encapsidation of a retroviral	
CC vector. The vectors can be used for gene therapy, production of a	
CC recombinant polypeptides or production of transgenic animals.	
SQ Sequence 578 BP; 94 A; 139 C; 172 G; 173 U;	

Query Match 100.0%; Score 127; DB 1; Length 578;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1	ugucgacagcgcccgccgcaaaugugugugcccaacacgagcgagcggaauaacuug 60
DB 452	UGUCGACAGCGCCCGCCGCAUAUUGUGUGCCACACCGCGCGGCUUGCGAUAUAUACUUG 511
QY 61	gagagucuuuuugccuacagugucuuuuguaucugucuccuuccuuccgagcg 120
DB 512	GAGAGUCUUUUUGCCUCCAGUGUCUUUGUACUCGUCUCCUCCUCCUCCGCGCG 571
QY 121	ggauggg 127
DB 572	GGAUGGG 578

RESULT 2
ID V63778 standard; DNA; 940 BP.
AC V63778;
DT 17-MAR-1999 (first entry)
DE Sequence derived from the 5' end of REV-A.
KW Reticuloendotheliosis virus type A; REV-A; 5' end genomic RNA;

KW type C retrovirus; internal ribosome entry site; IRES; encapsidation;
 KM retroviral vector; gene therapy; ss.
 OS Reticuloendotheliosis virus.
 PN FR2762615-A1.
 PD 30-OCT-1998.
 PF 28-APR-1997; 005203.
 PR 28-APR-1997; FR-005203.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PI Darlix JL, Gabus Darlix C, Lopez LM;
 DR WPI; 99-037487/04.
 PR Expression vectors containing IRES and/or encapsidation enhancer -
 PT derived from type C retrovirus other than FMLV and MCMV
 PS Claim 5; Page 31-32; 43pp; French.
 CC The present sequence is derived from the 5' end of the genomic RNA of
 CC Reticuloendotheliosis virus type A (REV-A). The specification also
 CC describes nucleotide sequence derived from all or part of the 5' end
 CC of the genomic RNA of a type C retrovirus other than Friend murine
 CC leukaemia virus (FMLV) and Moloney murine leukaemia virus (MCMV). The
 CC 5' derived sequences are used as an internal ribosome entry site (IRES)
 CC in a vector for permitting or enhancing the encapsidation of a retroviral
 CC vector. The vectors can be used for gene therapy, production of
 CC recombinant polypeptides or production of transgenic animals.
 CC Sequence 940 BP; 190 A; 225 C; 260 G; 265 U;

Query Match	100.0%;	Score 127;	DB 1;	Length 940;
Best Local Similarity	100.0%;	Pred. No. 2.7e-33;		
Matches 127; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

[illegible]

RESULT	3
x26113	
ID	x26113 standard; DNA; 4643 BP.

AC 20-MAY-1999 (first entry)
DE Fowlpox virus (FPV S) 5' LTR sequence.
KW Fowlpox virus; FPV; recombinant; reticuloendotheliosis virus; REV;
KM vaccine; chicken; LTR; long terminal repeat; ss.
OS Fowlpox virus.
PN MO9907852-A1.
PD 18-FEB-1999.
PE 07-AUG-1998; AU0628.
PR 08-AUG-1997; AU-008454.
PA (CSIR) COMMONWEALTH SCL & IND RES ORG.
PI Boyle DB, Coupar BEH, Gould AR, Hertig C;
DR WPI; 99-167428/14.
PT Recombinant vaccine against fowlpox virus - is free of contamination
PT from reticuloendotheliosis virus, used to prevent the virus
PT occurrence in chickens
PT
PS Claim 18; Fig 6; 127pp; English.
CC The invention relates to a recombinant fowlpox virus (FPV) whose nucleic
CC acid does not encode an active reticuloendotheliosis virus (REV), and
CC whose genome does not contain any REV sequence. The recombinant FPV is
CC used to produce a vaccine against fowlpox virus. The vaccine is used to
CC prevent the occurrence of fowlpox virus in chickens. The chickens can be
CC vaccinated at older than 1 day of age. Alternatively, chickens older than
CC 1 day can be vaccinated with the vaccine subsequent to vaccination with
CC FPV M (FPV mild strain) at 1 day of age. Prior vaccine and field strains
CC of FPV carry integrated REV sequences. Some of the FPV strains carry a
CC near full length provirus of REV, and can give rise to infectious REV
CC when transfected into cell cultures and when chickens are infected. The

CC present invention provides the means by which a FPV vaccine free from
CC contamination by REV may be produced. Sequences X26110-115 represent FPV
CC long terminal repeat sequences. The new recombinant FPV nucleic acid
CC molecule can comprise a FPV LTR sequence selected from the above.
SQ Sequence 4643 BP; 1326 A; 1015 C; 1127 G; 1175 T;

Query Match	98.7%;	Score 125.4;	DB 1;	Length 4643;
Best Local Similarity	70.9%;	Pred. No. 1.5e-32;		
Matches	90;	Conservative	36;	Mismatches 1;
				Indels 0;
				Gaps 0;

[illegible]

RESULT	4
Q76040	
Q76040 standard; DNA; 5519 BP.	

AC 07/06/94 (first entry)
 DT 20-JUL-1995
 DE Retrovirus vector pPol11-R3.
 KW Retrovirus; vector; pPol11-R3; spleen necrosis virus; SNV;
 KW cytomegalovirus; CMV; intermediate-early promoter; IE promoter;
 KW long terminal repeat; LTR; encapsidation; gene transfer;
 KW gene therapy; ss.
 OS Spleen necrosis virus.
 PN W09429437-A.
 PD 22-DEC-1994.
 PF 07-JUN-1994; U06415.
 PR 07-JUN-1993; US-073345.
 PA (UYNE-) UNIV NEW JERSEY.
 PI Dornburg RC;
 DR WPI; 95-036467/05.
 PT Recombinant retrovirus vector, contg. non-retroviral gene. - has
 PT ability to produce progeny virus, in helper cell which can infect
 PT host cell and form provirus
 PT Example; page 16-17; 28pp; English.
 PS New recombination-free, highly efficient retroviral vectors
 CC New recombination-free, highly efficient retroviral vectors
 CC pPol11-R1 (given in Q76038), pPol11-R2 (Q76039) and pPol11-R3
 CC (Q76040) were obtained by replacing the U3 region of the left
 CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV,
 CC and extension of the encapsidation region. The vectors allow
 CC cell-type specific gene expression and eliminate risks of
 CC downstream activation of cellular proto-oncogenes.
 CC Sequence 5519 BP; 1264 A; 1420 C; 1456 G; 1379 T;
 SQ

Query Match	91.88;	Score 116.6;	DB 1;	Length 5519;
Best Local Similarity	69.98;	Pred. No. 1.3e-29;		
Matches	86;	Conservative	33;	Mismatches 4;
				Indels 0;
				Gaps 0;

[illegible]

[illegible]

CC (Q76040) were obtained by replacing the U3 region of the left
CC LTR of spleen necrosis virus with the IE promoter/enhancer of CMV
CC and extension of the encapsidation region. The vectors allow
CC cell-type specific gene expression and eliminate risks of
CC downstream activation of cellular proto-oncogenes.
SQ Sequence 558 BP; 1265 A; 1423 C; 1457 G; 1503 T

RESULT 7
ID 076041 standard; DNA; 3878 BP.
AC 076041;
DT 20-JUL-1995 (first entry)
DE Retrovirus vector pPOL15-R1.
KW Retrovirus; vector; pPOL15-R1;
KW gene transfer; gene therapy; ss.
OS Spleen necrosis virus.
PN W09429437-A.
PD 22-DEC-1994.
PF 07-JUN-1994; U06415.
PR 07-JUN-1993; U5-073345.
PA (UYNE-) UNIV NEW JERSEY.
PI Dornburg RC;
DR WPI; 95-036467/05.
PT Recombinant retrovirus vector, contg. non-retroviral gene, - has
PT ability to produce progeny virus, in helper cell which can infect
PT host cell and form provirus
PS Example; Page 17; 28pp; English.
CC The universal retroviral vector pPOL15-R1 (given in 076041) was
CC obtained by replacing the SV40 promoter and hpt gene of pPOL11-R1
CC (Q76038) with the multiple cloning site of pBluescript II KS. The
CC vectors allow cell-type specific gene expression and eliminate risks
CC of downstream activation of cellular proto-oncogenes.
SQ Sequence 3878 BP; 922 A; 948 C; 988 G; 1020 T;

```

Query Match          89.4%; Score 113.6; DB 1; Length 3878;
Best Local Similarity 69.2%; Pred. No. 1.1e-28;
Matches 83; Conservative 33; Mismatches 4; Indels 0; Gaps 0

QY 1 ugucgacacgcgccucgcgaauugugugccacacgcgcgcgcguugcgaaauacuug 60
   |:||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1153 TGTGACAGCGCCTTGCGAATTGGTGTACCCACACCGCGCGCTGCGAATAATACTTTG 1212
      gagagucuuuugcucuccagugucucucguuuugacuugucucucucuccucucgcgcg 120
      ||||| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 1213 GAGAGCCTTTTGCCCTCCAGTGTCTTCGCTGTACTCGTCCCTCTCCCTCCGCGCG 1272

RESULT 8
ID Q80523/c
AC Q80523;
DT 15-JUL-1995 (first entry)
DE Human mcl-1 gene.
KW Myeloid cell leukemia-associated gene; mcl-1; diagnostic;

```


Tue Aug 8 08:33:22 2000

us-09-214-124-2_copy_452_578.rng

THIS PAGE BLANK (USPTO)